Planetary and Space Science

Volume 8 1961

Editor-in-Chief: MILTON GREENBERG

Regional Editors:

D. R. BATES
W. W. KELLOGG
D. F. MARTYN
T. NAGATA
W. ORR ROBERTS
H. FRIEDMAN
A. I. LEBEDINSKY
S. K. MITRA
M. NICOLET
F. L. WHIPPLE

PERGAMON PRESS

NEW YORK · OXFORD · LONDON · PARIS

Planetary and Space Science

Editor-in-Chief MILTON GREENBERG

REGIONAL EDITORS

D. R. BATES (United Kingdom)
A. I. LEBEDINSKY (Soviet Union)
S. K. MITRA (Far East)
M. NICOLLT (Europe)
W. ORR ROBERTS (North America)
H. FRIEDMAN (North America)
W. W. KELLOGG (North America)
D. F. MARTYN (Southern Hemisphere)
T. NAGATA (Japan)
F. L. WHIPPLE (North America)

EDITORIAL ADVISORY BOARD

H. O. G. ALFVEN—Stockholm
Sir EDWARD APPLETON—Edinburgh
H. W. BABCOCK—Pasadena
J. BARTELS—Gottingen
FR. BECKER—Bonn
W. J. G. BEYNON—Aberystwyth
M. BOSSOLASCO—Genova
R. L. F. BOYD—London
R. N. BRACEWELL—Stanford
Sir EDWARD BULLARD—Cambridge, England
S. CHANDRASEKHAR—Williams Bay
S. CHAPMAN—College, Alaska
J. COULOMB—Paris
B. W. CURRIE—Saskatoon
H. ELLIOT—London
C. T. ELVEY—College, Alaska
I. ESCOBAR—La Paz
H. FLOHN—Offenbach
L. GOLDBERG—Ann Arbor
F. GRAHAM SMITH—Cambridge
W. GROTH—Bonn
S. H. HALL—London
L. HARANG—Oslo
T. HATANAKA—Tokyo
B. HAURWITZ—New York City
A. R. HIBBS—California
R. E. HOLZER—Los Angeles
J. T. HOUGHTON—Oxford
H. C. VAN DE HULST—Leiden
L. G. H. HUXLEY—Adelaide
R. JASTROW—New York City
R. C. JENNISON—Jodrell Bank
J. KAPLAN—Los Angles
A. KIMPARA—Toyokawa-City
V. I. KRASOVSKII—Moscow
E. A. KREIKEN—Ankara
G. P. KUIPER—Williams Bay
N. LAFFINEUR—Paris
H. E. LANDSBERG—Washington, D.C.
V. LAURSEN—Charlottenlund
Sir BERNARD LOVELL—Jodrell Bank
T. F. MALONE—Hartford
K. MAEDA—Kyoto
Sir HARRIE MASSEY—London
A. MAXWELL—Fort Davis
E. F. MCCLAIN—Washington
P. M. MILLMAN—Ottawa
F. MÖLLER—Mainz
M. G. MORGAN—Hanover, New Hampshire
H. NORINDER—Uppsala
R. W. PORTER—Schenectady
K. R. RAMANATHAN—Ahmedabad
G. RHIGHINI—Florence
J. SAYERS—Birmingham
B. F. J. SCHONLAND—Harwell
L. I. SEDOV—Moscow
A. H. SHAPLEY—Boulder
A. F. SPILHAUS—Minneapolis
E. A. STEWARDSON—Leicester
O. STRUVE—Green Bank, West Virginia
R. C. SUTCLIFFE—London
P. SWINGS—Liege
A. T. VASSY—Paris
L. VEGARD—Oslo
K. WATANABE—Honolulu
H. C. WEBSTER—Brisbane
H. WEXLER—Washington, D.C.
J. T. WILSON—Toronto

Publishing Offices: Pergamon Press Ltd., Headington Hill Hall, Oxford. (Oxford 64881)
Annual Subscription (including postage): (A) for libraries—£21 (U.S.A. $60.00).
(B) for subscribers certifying that the journal is for their own personal use—£5 5s. ($15.00).

PERGAMON PRESS, INC.

122 EAST 55TH STREET, NEW YORK 22, NY. 4 & 5 FITZROY SQUARE, LONDON, W.1

PRINTED IN GREAT BRITAIN BY THE LEWES PRESS WIGHTMAN AND CO. LTD., LEWES, SUSSEX.

CONTENTS

NUMBER 1

G. FIELDER: Contraction and expansion of the Moon 1
S. N. GHOSH and K. D. SHARMA: The number of N_2 molecules in Earth's atmosphere from rocket data . 9
C. J. LOUGHNAN: Longitudinal dependence of radiation-belt scattering, and primary auroral particles . . 13
L. D. KAPLAN: A new interpretation of the structure and CO_2 content of the Venus atmosphere . . 23
E. E. BRUESCHKE, R. H. SUESS and M. WILLARD: The viability of micro-organisms in ultra-high vacuum . 30
J. A. REES: Diffusion coefficients determined from sodium vapour trails 35
D. E. SMITH: Determination of the earth's gravitational potential from satellite orbits 43

NUMBER 2

M. H. REES and C. S. DEEHR: The aurora of 27 November 1959 at College, Alaska, including observations of a high altitude red arc 49
M. H. REES: Excitation of high altitude red auroral arc 59
J. HUGILL: The measurements of random radio frequency fields by the use of short electric and magnetic dipoles 68
T. N. L. PATTERSON: Whistler dispersions for a model of the ionic distribution in the exosphere . . . 71

SELECTED PAPERS AND ABSTRACTS TRANSLATED FROM *ARTIFICIAL EARTH SATELLITES U.S.S.R.*
A. I. LUR'E: The equations of disturbed motion in the Kepler problem 79
T. N. NAZAROVA: Investigation of meteoric particles on the third Soviet artificial earth satellite . . 82
M. D. KISLIK: The path of an artificial earth satellite in the normal gravitational field of the earth . . 86
B. V. RAUSHENBAKH and E. N. TOKAR': Problems of guidance in interplanetary space 97
L. I. SEDOV: Orbits of space rockets in the direction of the Moon 107
ABSTRACTS 117

RESEARCH NOTES
K. A. POUNDS: The soft X-ray absorption of aluminium and rocket measurements of solar radiation . 120
A. R. HIBBS: Linear arrays of small craters 121
BOOK REVIEWS 122

NUMBER 3/4

V. J. KRASOVSKII: Some results of investigations of aurorae and night airglow during the I.G.Y. and I.G.C. 125
B. HULTQVIST: Circular symmetry in the geomagnetic equatorial plane for auroral phenomena . . . 142
J. D. WHITEHEAD: The absorption of radio waves in the ionosphere with respect to geomagnetic current control and of the proposed models—A critical survey 151

SELECTED PAPERS AND ABSTRACTS TRANSLATED FROM *ARTIFICIAL EARTH SATELLITES U.S.S.R.*
I. M. YATSUNSKII: Determination of illumination conditions and periods of an artificial satellite in shadow and in sunlight 159
YU. G. SHAFER and A. V. YARYGIN: Measurement of cosmic rays on geophysical rockets 165
V. A. SARYCHEV: Influence of the earth's oblateness on the rotational motion of an artificial satellite . . 173
V. G. ISTOMIN: Variation in the positive ion concentration with altitude from data of mass spectrometry measurements on the third satellite 179
K. I. GRINGAUZ and V. A. RUDAKOV: Measurements of electron concentration in the ionosphere up to altitudes of 420-470 km, carried out during the I.G.Y. by means of radio-wave transmissions from the geophysical rockets of the Academy of Sciences U.S.S.R. 183
ABSTRACTS 194

RESEARCH NOTES
V. I. KRASOVSKII: On the night sky hydroxyl emission 197
M. H. REES: On the excitation of subvisual OI λ λ 6300-6364 arcs at middle latitudes 198
E. Y. C. INN: Further comments on the origin of the *D*-region 200

AUTHOR INDEX TO VOLUME 8

BRUESCHKE, E. E. 30

DEEHR, C. S. 49

FIELDER, G. 1

GHOSH, S. N. 9
GRINGAUZ, K. I. 183

HIBBS, A. R. 121 (R)
HUGILL, J. 68
HULTQVIST, B. 142

INN, E. Y. C. 200 (R)
ISTOMIN, V. G. 179

KAPLAN, L. D. 23
KISLIK, M. D. 86
KRASOVSKII, V. I. 125, 197 (R)

LOUGHNAN, C. J. 13
LUR'E, A. E. 79

NAZAROVA, T. N. 82

PATTERSON, T. N. L. 97
POUNDS, K. A. 120 (R)

RAUSHENBAKH, B. V. 97
REES, J. A. 35
REES, M. H. 49, 59, 198 (R)
RUDAKOV, V. A. 183

SARYCHEV, V. A. 173
SEDOV, L. I. 107
SHAFER, YU. G. 165
SHARMA, K. D. 9
SMITH, D. E. 43
SUESS, R. H. 30

TOKAR', E. N. 97

WHITEHEAD, J. D. 151
WILLARD, M. 30

YARYGIN, A. V. 165
YATSUNSKII, I. M. 159

PLANETARY AND SPACE SCIENCE

Publisher's notice to contributors and readers about the time taken for the publication of accepted contributions to the Journal.

Positive arrangements have now been made with the printer ensuring regular and rapid appearance of issues of the Journal which enables us to guarantee to contributors that papers submitted and accepted by the Board of Editors for publication in *Planetary and Space Science* will be published within 90–130 days depending on the speed that each author returns galleys approved for press.

PERGAMON PRESS LIMITED.

Planetary and Space Science

CUMULATIVE INDEXES

Volumes 6–8
1961

Editor-in-Chief: MILTON GREENBERG

Regional Editors:

D. R. BATES
W. W. KELLOGG
D. F. MARTYN
T. NAGATA
W. ORR ROBERTS
H. FRIEDMAN
A. I. LEBEDINSKY
S. K. MITRA
M. NICOLET
F. L. WHIPPLE

PERGAMON PRESS

NEW YORK · OXFORD · LONDON · PARIS

Subject Index

Absorption
- microwave radiation in a plasma, **6**, 105.
- radiowaves in ionosphere: **8**, 151.
- soft X-rays by aluminium: **8**, 169

Air, high temperature
- electron concentration in closed form: **6**, 79.
- collision frequency: **6**, 24.
- cross-sections for constituents: **6**, 24

Air, seeded: **6**, 79, 172.
Airglow, artificial, measurement of diffusion coefficient: **8**, 35.
Antennae
- ambipolar diffusion, effect on breakdown: **6**, 149
- power handling at high altitudes: **6**, 142
- radiation from slotted-sphere: **6**, 116
- radomes, microwave breakdown at high altitudes: **6**, 155.
- R.F. breakdown in presence of sheath: **6**, 149.

Arcing, at high altitudes: **7**, 11
Atmosphere of Earth
- abundance of N_2, **8**, 9.
- ARDC 1959 model, review: **7**, 454.
- model for downward transfer of energy: **7**, 434.
- thermal structure of ozonosphere from satellite measurements: **7**, 423.

Aurora
- circular symmetry in geomagnetic equatorial plane: **8**, 142.
- excitation of high altitude red arcs: **8**, 59.
- excitation of subvisual OI arcs at middle latitudes: **8**, 198.
- observations during I.G.Y. and I.G.C.: **8**, 125.
- observations in Alaska including high altitude red arcs: **8**, 49.
- precipitation of trapped particles: **8**, 13.
- transmission of UHF signals through: **7**, 272.

Book reviews
- *An introduction to celestial mechanics* by T. E. Sterne: **8**, 122.
- *Cosmic radio waves* by I. S. Shklovskii, trans. by R. B. Rodman and C. M. Varsavsky: **8**, 122.
- *Physics of the solar chromosphere* by Richard N. Thomas and R. Grant Athay: **8**, 122.

Carbon dioxide, in atmosphere of Venus: **8**, 23.
Collision frequencies, high temperature air: **6**, 24.
Communications, by satellites: **7**, 254.
Computers
- high speed memory element: **6**, 184.
- random failures in components: **7**, 242
- simulation of missiles: **7**, 179.

Cosmic rays, measurement on geophysical rockets: **8**, 165.

Diffusion
- ambipolar, effect on antenna breakdown: **6**, 149.
- coefficients, determined from sodium vapour trails: **8**, 35.

Electron concentration in high temperature air: **6**, 79.
Exosphere, model of ion distributions in: **8**, 71.

Guidance
application of dynamic programming: **7**, 176.
control of soft landings on the Moon: **7**, 70.
high precision stellar navigator: **7**, 57
homing: **7**, 125
inertial components to supplement astro-fixes: **7**, 19.
interplanetary and lunar vehicles: **7**, 64, 94.
miss distances for lunar and interplanetary trajectories: **7**, 125.
orbital transfer for satellites: **7**, 132.
orbits of space rockets in the direction of the Moon: **8**, 107.
problems in interplanetary space: **8**, 97.
reconnaissance missions to Venus and Mars: **7**, 40.
static inverter: **7**, 1.
trajectories for continuous propulsive thrust: **7**, 145.
vertical azimuth transfer system: **7**, 108.

Hydroxyl emission, nightglow: **8**, 197.
Hypersonics
collision frequencies in high temperature air: **6**, 24.
effect of errors in rate constants on calculated shock layer electron densities: **6**, 74.
electron concentration in closed form for high temperature air: **6**, 79.
ionized shock layer for blunt shapes: **6**, 10.
plasma data of equilibrium air: **6**, 47.
radio signal interference due to ionized layer: **6**, 1; **7**, 286.
rate of ionization behind shock waves: **6**, 94.

Ionization
non-equilibrium, in shock waves: **6**, 74.
rate behind shock waves in air: **6**, 94.
rate in lower atmosphere: **8**, 9.
Ionosphere
D-region, origin of: **8**, 200.
electron concentration: **8**, 183.
positive ion concentration from Sputnik III: **8**, 179
Ions, in exosphere: **8**, 71.

Kepler problem, equations of disturbed motion: **8**, 79.

Micrometeorites
flux measurement: **7**, 417.
on Sputnik III: **8**, 82.
Micro-organisms, viability in ultra-high vacuum: **8**, 30.
Microwaves
radiation, absorption in plasma: **6**, 105.
techniques for hypersonic ballistic ranges: **6**, 186.
Missiles
damping of structural bending: **7**, 26.
performance of static inverter: **7**, 1.
radar tracking: **7**, 81.
reliability: **7**, 208, 217, 230.
simulation by a digital computer: **7**, 179.
trajectory charts: **7**, 118.
Moon
expansion and contraction: **8**, 1.

linear array of small craters: **8**, 121
orbits of rockets: **8**, 107.
soft landings on: **7**, 70.

Nightglow, hydroxyl emission: **8**, 197.
Nitrogen, molecular, abundance in terrestrial atmosphere: **8**, 9.

Plasma
coated sphere, echo area: **6**, 133.
data of equilibrium air: **6**, 47.
from seeded cyanogen flame, propagation in: **6**, 172.
non-uniform, interaction with microwave radiation: **6**, 105.
propagation properties: **6**, 1.
sheath, on spherical satellite, radar cross-section: **6**, 123.
sheath, physical properties around projectiles: **6**, 196.
sheath, propagation through, in hypersonic low density wind tunnel: **6**, 180.
waves, attenuation of: **6**, 1.
waves, in a magnetic field: **6**, 100.
WKB approximation in: **6**, 105.

Radiation, solar, rocket measurements: **8**, 120.
Reconnaissance, mission to Venus and Mars: **7**, 40.
Re-entry
effects of impact: **7**, 386.
radiation from slotted-sphere antenna with plasma sheath: **6**, 116.
radio signal interference due to ionized layer: **6**, 1; **7**, 286.
telecommunications during: **6**, 207.
R.F. fields, random, measurement using short electric and magnetic dipoles: **8**, 68.

Satellites
bi-elliptical transfer between coplanar circular orbits: **7**, 164.
communications by means of: **7**, 254.
determination of Earth's gravitational potential: **8**, 43.
effect of Earth's oblateness on rotational motion: **8**, 173.
Hohmann transfer: **7**, 132, 164.
illumination conditions and periods in shadow and sunlight: **8**, 159.
impact of meteoric particles on Sputnik III: **8**, 82.
orbital transfer: **7**, 132.
orbits derived from gravitational model of the Earth: **7**, 154.
orientation effects: **7**, 445.
path in normal gravitational field of the Earth: **8**, 86.
positive ion concentrations from Sputnik III: **8**, 173.
radar cross-section: **6**, 123.
Shock layer
blunt shapes: **6**, 10.
non-equilibrium ionization in: **6**, 74.
transmission of e.m. radiation through **6**, 1.
Sodium, vapour trails: **8**, 35.
Solar, radiation, rocket measurements: **8**, 120.
Space flight
bioastronautics research: **7**, 356.
closed cycle life support system: **7**, 370.
ecology problems: **7**, 345.
effects of impact: **7**, 386.
hypothetical orbital flight: **7**, 309.

man in space: **7**, 301.
micro-organisms: **8**, 30.
physiological effects: **7**, 378, 395.
space simulators: **7**, 324, 335.
Thermal impact: **7**, 427.
Tracking
generalized trajectory charts: **7**, 118.
high precision 200 million mile system: **7**, 94.
photographic satellite position determination: **7**, 76.
radar, for satellites and missiles: **7**, 81.

Van Allen radiation
composition of lower belt: **7**, 411.
longitudinal dependence: **8**, 13.
Velocity, corrections for space vehicle: **7**, 40.
Venus, CO_2 content of atmosphere: **8**, 23.

Waves, plasma, in magnetic field: **6**, 100.
Whistlers, dispersions for an ionospheric model: **8**, 71.
W.K.B. approximation, in a plasma: **6**, 105.

Author Index

(B signifies book review and R research note)

Allis, W. P. **6**, 100
Altman, S. P. **7**, 132
Anderson, G. P. **7**, 242
Arthur, J. S. **7**, 81

Bachynski, M. P. **6**, 24
Baldwin, K. M. **6**, 207
Barker, C. L., Jnr. **7**, 335
Basset, O. E. **6**, 207
Battin, R. H. **7**, 40
Bellman, R. **7**, 176
Berkofsky, L. **7**, 434
Bock, C. D. **7**, 57
Bowring, J. I. R. **7**, 309
Bracha, V. J. **7**, 208
Brady, J. S. **7**, 301
Brechner, V. O. **7**, 378
Brown, R. Hanbury, **8**, 123 (B)
Brueckner, K. A. **6**, 105
Brueschke, E. E. **8**, 30

Champion, K. S. W. **7**, 454
Chuan, R. L. **6**, 180
Cohen, A. D. **7**, 118
Cohen, H. A. **7**, 417
Cook, G. E. **8**, 122 (B)
Crooks, J. W. **7**, 94
Cutler, C. C. **7**, 254

Darlington, S. **7**, 70
Deehr, C. S. **8**, 49
Dieter, N. H. **7**, 76
Dolph, C. L. **6**, 123
Dreyfus, S. **7**, 176

Ebert, B. P. **7**, 309
Emmerich, C. L. **7**, 108
Epstein, A. H. **7**, 179

Fielder, G. **8**, 1
Fiskin, J. M. **6**, 47
Freed, L. E. **7**, 26

Ghosh, S. N. **8**, 9
Green, R. B. **6**, 133
Greyber, H. D. **6**, 105
Gringauz, K. I. **8**, 183

Hawthorne, E. I. **6**, 207
Henry, J. C. **7**, 81
Hibbs, A. R. **8**, 121 (R)
Hochstim, A. R. **6**, 79
Hoelker, R. F. **7**, 164
Holly, F. E. **7**, 411
Huber, P. W. **6**, 172
Hughes, J. V. **7**, 145
Hugill, J. **8**, 68
Hultqvist, B. **8**, 142

Ingalls, R. P. **7**, 272
Inn, E. C. Y. **8**, 201 (R)
Istomin, V. G. **8**, 179

James, J. C. **7**, 272
Johnson, R. G. **7**, 411
Johnston, T. W. **6**, 24

Kalaba, R. **7**, 176
Kaplan, L. D. **8**, 23
Karmiol, E. D. **7**, 230
Keenan, P. P. **6**, 155
King, J. I. F. **6**, 105; **7**, 423
Kislik, K. D. **8**, 86
Kizner, W. **7**, 125
Klein, M. M. **6**, 105
Konikoff, J. J. **7**, 370
Kornhauser, M. **7**, 386
Krasovskii, V. I. **8**, 125, 197 (R)

Langberg, E. **6**, 207
Laning, J. H. **7**, 40
Lawton, R. W. **7**, 386
Lin, S. C. **6**, 94
Lipow, M. **7**, 217
Lloyd, D. K. **7**, 217
Loughnan, C. J. **8**, 13
Lur'e, A. I. **8**, 79
Luton, W. B. **7**, 324

Macdonald, A. D. **6**, 149
McNally, D. **8**, 122 (B)
Maiden, C. J. **6**, 196
Marini, J. W. **6**, 116
Michaels, J. E. **7**, 427
Miczaika, G. R. **7**, 76
Miller, K. L. **7**, 26
Minzner, R. A. **7**, 454
Morita, T. **6**, 142
Mundo, C. J. **7**, 19

Nauman, R. J. **7**, 445
Nazarova, T. N. **8**, 82
Nelson, J. R. **7**, 345
Nomicos, G. N. **7**, 145

Papa, R. J. **6**, 100
Patterson, T. N. L. **8**, 71
Peters, L. **6**, 133
Peterson, L. E. **7**, 242
Peterson, N. W. **7**, 179
Pettengill, G. H. **7**, 81
Pittman, C. W. **7**, 64
Pogrund, R. S. **7**, 395
Pond, H. L. **7**, 454
Pounds, K. A. **8**, 120 (R)
Primich, R. I. **6**, 186
Quinn, F. B. **7**, 378

Raushenbakh, B. V. **8**, 97
Rees, J. A. **8**, 35
Rees, M. H. **8**, 49, 59, 198 (R)
Ridyard, H. W. **6**, 10
Rudakov, V. A. **8**, 183

St. Pierre, C. **6**, 196
Sarychev, V. A. **8**, 173
Savet, P. H. **7**, 154
Scharfman, W. E. **6**, 142
Schroeder, W. **7**, 64
Sebring, P. B. **7**, 81
Sedov, L. I. **8**, 107
Shafer Yu. G. **8**, 165
Shapiro, R. **7**, 434
Sharma, K. D. **8**, 9
Shkarovskii, I. P. **6**, 24
Silber, R. **7**, 164
Sisco, W. B. **6**, 47
Smith, D. E. **8**, 43
Spink, B. R. **7**, 11
Steffe, W. **7**, 1
Stone, M. L. **7**, 272
Suess, R. H. **8**, 30

Taylor, W. C. **6**, 1; **7**, 286
Thompson, A. B. **7**, 324
Tokar, E. N. **8**, 97

Wal, L. Van Der **7**, 356
Wanlass, C. L. **7**, 184
Wanlass, S. D. **7**, 184
Weil, H. **6**, 123
Whitehead, J. D. **8**, 151
Whitmer, R. F. **6**, 149
Willard, M. **8**, 30
Wisnieff, R. E. **7**, 108

Yarygin, A. V. **8**, 165
Yatsunskii, I. M. **8**, 159
Youtcheff, J. S. **7**, 230

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 1–8. Printed in Great Britain.

THE CONTRACTION AND EXPANSION OF THE MOON

GILBERT FIELDER*
University of London Observatory, Mill Hill Park, London N.W.7

(*Received* 17 *April* 1961)

Abstract—Measurements of craters near to the centre of the Moon's disk are used to show that the craters are distorted preferentially, their longer axes lying parallel to the most prominent family of the grid system. It is shown that, on average, the percentage distortions are greater for the older craters, and that the large craters are generally older than the small craters which we now see; but old, small craters would be effaced by crustal movements. At any given time, craters of all sizes were possibly being produced. Taken together with other evidence for a recent tensional phase in the Moon's history, the present results lead to the conclusion that the crustal stresses which distorted the craters by thrust-faulting decreased in magnitude as time elapsed, and then reversed in sign.

It is proposed that, in the early history of the Moon, a general protuberance of the body of the Moon in the direction of the Earth tended to settle, gradually, to a form corresponding to a state of minimum gravitational potential. It is also suggested that, during this time, radioactivity was heating the Moon and that heat accumulated most rapidly in the innermost parts of the Moon, so that the volumetric expansion of the core exceeded that of the crust. As time elapsed, the crust itself was, therefore, subject to two opposing forces. The crustal pressures were reduced, the crust passed through a state of zero stress, and the stresses reversed in sign, to become tensions.

1. EVIDENCE FOR COMPRESSION OF THE MOON'S CRUST

Various phenomena in the Ariadaeus-Hyginus regions of the Moon (Fig. 1) may be explained[1] on the hypothesis that there were, at one time, uniform thrusts in the Moon's crust.† One difficulty of this hypothesis is that the Ariadaeus and Hyginus Rilles are supposed to have originated in wrench fractures in which dykes formed; and it is known from geological studies that the magmas of dykes are seldom intruded along wrench faults because the residue of the thrusts which formed them usually prevents the fractures from opening. In order to overcome this difficulty, the writer suggested that there was a tensile phase in the Moon's crustal history which post-dated the compressive phase and, hence, opened the fractures. Evidence in support of this suggestion has now been found. (See Section 3, below).

Von Bulow[2,3] prefers to explain the origin of the rilles in this region of the Moon by postulating a widespread updoming which gave rise to tensions in the surface. This hypothesis would not, by itself, explain the particular relative orientations of the major rilles and ridges in the Ariadaeus-Hyginus regions. It is also important to consider whether the major-distortions of the many evidently very eroded craters in this area can be explained at all without resorting to a surface compressive phase.

The craters in the area delineated by the Mount Wilson plate C4-b of Kuiper's *Photographic Lunar Atlas*[4], which depicts the Ariadaeus-Hyginus regions, have now been measured in order to obtain quantitative data on the degrees of distortion which they display. Of the craters identified in Fig. 1, 134 were found to be suitable for measurement. Now the most

* Imperial Chemical Industries Research Fellow of the University of London, U.L. Observatory, Mill Hill Park, London, N.W.7.

† In this discussion, frequent mention will be made of the "crust" of the Moon. There is no evidence that the Moon has a crust which is in any way analogous to that of the Earth, but the concept of a lunar "crust" and "core" will be used, here, to distinguish some shell of rock which is closest to the surface from the remainder of the Moon.

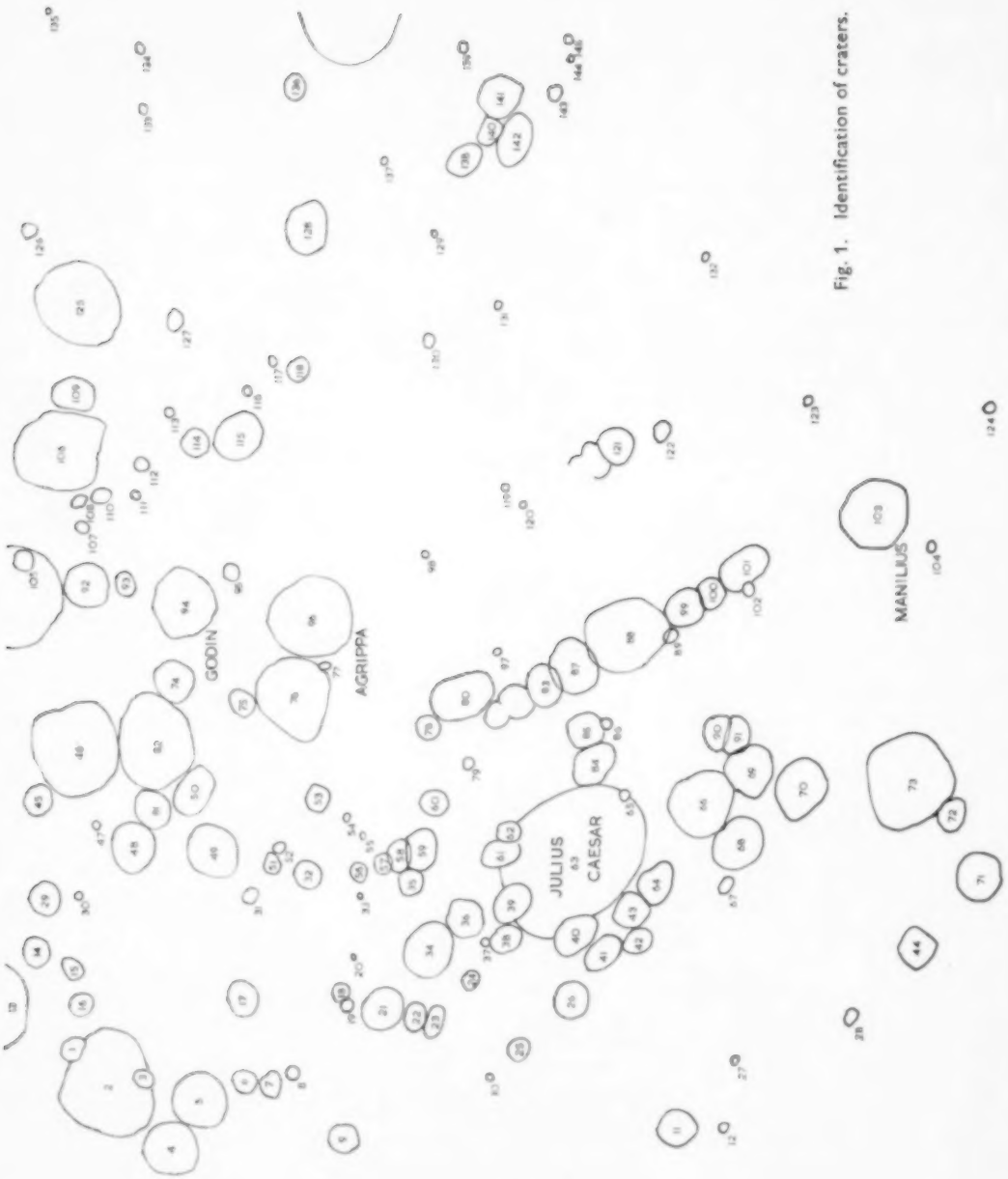


Fig. 1. Identification of craters.

prominent family of the grid system[7] in this vicinity of the Moon runs, on average, roughly parallel to the limb, and there is a less well-developed family of distortions which runs orthogonally to it. Measurements y and x (cf. Fig. 2), the orthogonal axes of craters, were made in the directions of these two most prominent families, respectively, and no restrictions were imposed on the selection of craters by virtue, for example, of their size or degree of erosion, except in so far as that no attempt was made to measure craters $\gtrsim$ 3 km in diameter. The measurements were then corrected for curvature, account being taken of the librations.

In Fig. 2, R represents the original radius of a crater and d and c the displacements of its walls

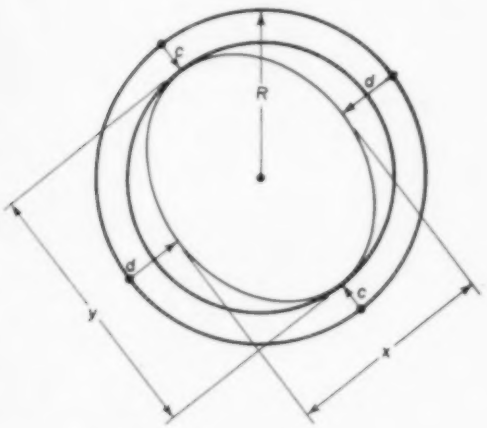


Fig. 2. The distortion of a crater.

along the minor and major semi-axes, respectively. It is seen that

$$\frac{d}{R} = 1 - \frac{x}{y+2c}. \qquad (1)$$

The quantity c is unknown: it is not measureable, but is likely to be small in comparison with y. Even if c is not very small, the minimum percentage distortion (which will be referred to, simply, as the percentage distortion) is given by

$$\Pi = \left(\frac{d}{R}\right)_{\min} . 100 = \left(1 - \frac{x}{y}\right) . 100\%. \qquad (2)$$

A plot (Fig. 3) of Π against y (which may be regarded as a measure of the size of a crater, but is actually the lower limit of the original diameter, as may be seen from Fig. 2) shows that, on average, the percentage distortion of a crater does not vary significantly with its size. Points which lie on the abscissa $\Pi = 0$, in Fig. 3, correspond to circular craters. Of the 134 points, 83 per cent of them lie above this line. This proves that the craters are distorted preferentially, with their longer axes in the direction of the most prominent family of the grid system.

In Table 1, the craters have been split into three size-groups y_1, y_2, y_3, defined in units of cm of plate, where 1 cm $\cong$ 13·7 km. The relative age of a crater was decided rather arbitrarily, but quite decisively, by combining characteristics such as ease of recognition of a crater and the height, or degree of erosion, of its walls; but *not* by considering the degree of distortion of its walls. Segregation of the craters (Table 2) in accordance with their age shows, that, in each of the size groups defined above, the mean percentage distortion of the old craters is considerably greater than that of the young craters. On average, $\Pi = 16$ per cent for the old craters and only 7 per cent for the young craters. This proves that the forces responsible for the deformations acted on the older craters for longer times.

The actual numbers of old and young craters are given in Table 3. It was found that 71 per cent of the craters in the y_1 group were to be identified with young craters but that only 27 per cent and 22 per cent of the craters in the groups y_2 and y_3, respectively, were young. Hence, in general, it is seen that a small crater is younger than a larger crater.

Again, whilst most craters show a "positive" distortion ($y - x > 0$), some show a "negative" distortion ($y - x < 0$), and 83 per cent of these are to be identified with the above-mentioned young craters. Of the 23 craters showing negative distortion, 83 per cent are found in the y_1 group and only about 9 per cent in each of the other size groups. The fact that some craters do show negative distortion may be explained, for example, by distortional scatter.

Table 1.

Group	Size range (cm plate: 1 cm ≡ 13·7 km)	Identification of Craters (cf. Fig. 1)		
		Old	Young	Negatively distorted
y_1	0–1·49	1, 14, 22, 23, 24, 35, 38, 45, 52, 57, 58, 60, 67, 72, 75, 89, 100, 107, 108, 110, 112, 114, 118, 126, 138, 140	3, 6, 7, 8, 16 17, 18, 19, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 37, 41, 42, 44, 47, 51, 53, 54, 55, 56, 61, 62, 65, 77, 78, 79, 86, 93, 95, 97, 98, 102, 104, 105, 111, 113, 116, 117, 119, 120, 122, 123, 124, 127, 129, 130, 131, 132, 133, 134, 135, 136, 137, 139, 143, 144, 145	3, 6, 17, 33, 44, 51, 55, 56, 75, 93, 100, 102, 107, 111, 122, 133, 135, 139, 145
y_2	1·50–2·99	34, 36, 39, 40, 48, 50, 59, 64, 66, 68, 70, 81, 83, 84, 85, 87, 90, 91, 92, 99, 101, 109, 115, 142	5, 43, 49, 71, 74, 94, 103, 128, 141	49, 94
y_3	3·00–4·5	46, 76, 80, 82, 88, 106	96, 125	73, 125

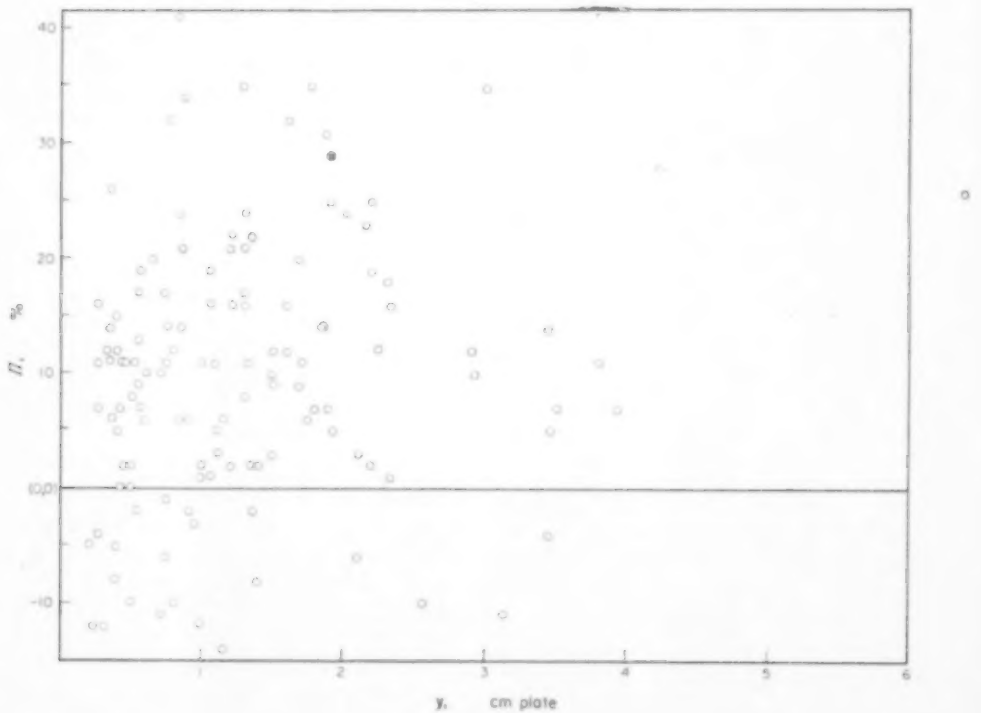


Fig. 3. The percentage distortions of craters in the Hyginus region.

Table 2.

Group	Mean percentage distortion		Mean actual distortion (cm plate)	
	Old	Young	Old	Young
y_1	16·7	6·1	0·12	0·05
y_2	16·9	4·7	0·38	0·11
y_3	14·6	1·9	0·55	0·07

Table 3.

Group	Total number of craters	Number of old craters	Number of young craters	No. of neg. distorted craters	Percentage of young craters
y_1	91	26	65	19	71
y_2	33	24	9	2	27
y_3	9	6	2	2	22

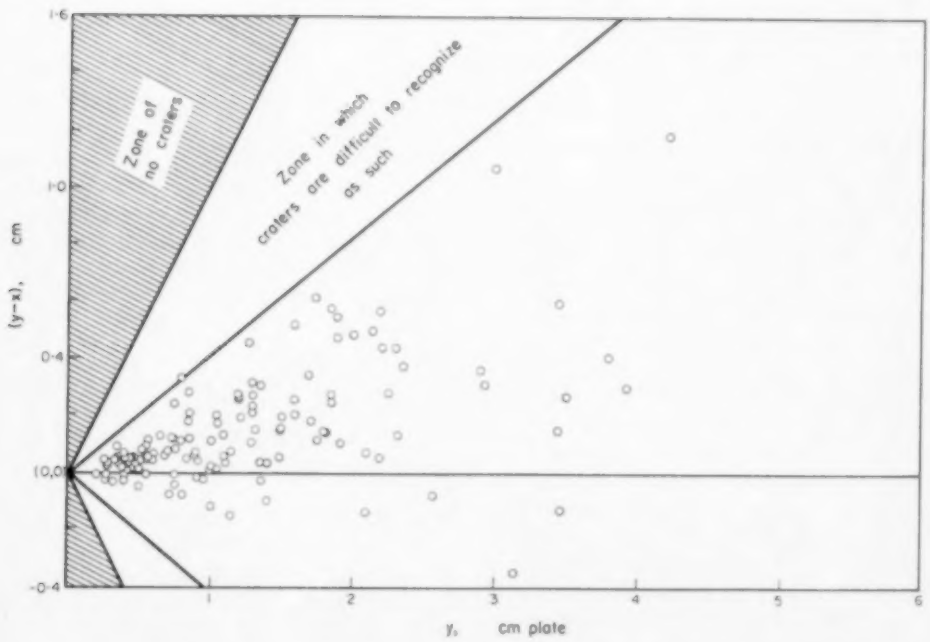


Fig. 4. The actual distortions of craters in the Hyginus region.

The quantity $y-x$ is a measure of the minimum value of the actual distortion of a crater, and this has been plotted (Fig. 4) against y. Because small craters which were formed earlier than a certain epoch in the Moon's history would be completely effaced by crustal movements (see the zone of no craters in Fig. 4, bounded by the line $x=0$) it is possible that craters of all sizes were being produced at all times, there being, at any given time, more small craters than large. Whether this is so, or whether there were relatively fewer small craters in the past, in no way affects the argument, confirmed by the above tests, that the small craters which we see at the *present* time, are, in general, the youngest.

These results lead to the general conclusion that the largest craters in the area in question were formed first, and that either

(i) the crustal stresses which distorted the craters decreased in magnitude as time elapsed, or

(ii) the crustal stresses were more or less uniform in magnitude throughout lunar history.

In the opinion of the writer, the results can only be explained on the hypothesis of crustal overthrusting.

2. A TENSILE PHASE IN LUNAR HISTORY

Observing from the Pic-du-Midi, the writer[5] studied the fine array of ridges which radiate in a north-easterly direction from Aristillus. On photographs, these ridges look like thin, straight rays; however, they were, in fact, observed to cast shadows. It was quite impracticable to think of these radiating spines and valleys as having been formed by blocks of matter ejected horizontally, during the formation of Aristillus, to plough grooves in the Moon's surface, for the ridges were too straight and uniform in width. The writer had the impression that this was a system of rift valleys caused by block-faulting. In theory, such faults could be produced by lateral thrusts *or* tensions. Several lunar craters exhibit these phenomena.

A theoretical approach to the problem of the nature of the Moon's crustal stresses, based on measurements of radiating ridge-systems of this type, led B. Warner[6] and the present writer to conclude that, when certain craters were formed, the surface layers of the Moon were in *tension*. This apparently unambiguous result was unexpected, because it appeared to contradict the results mentioned above.

3. RELATIVE AGES OF SOME LUNAR FORMATIONS

The few well-formed large craters in Fig. 1, such as Godin, Agrippa, and Manilius, are evidently to be considered as younger than the greatly eroded ghost craters in the same area. Now the characteristic patterns of radiating ridges which are associated with craters formed in a medium in tension have been detected by Warner[6] around the following craters:—

Agrippa, Archimedes, Aristillus, Aristoteles, Autolycus, Bullialdus, Burg, Copernicus, Davy, Eratosthenes, Harpalus, Hercules, Herschel, Lambert, Plinius, Thebit, Theophilus, Timocharis.

Without exception, these craters have comparatively cleanly-sculptured ringwalls. All, except for Archimedes, have, in addition, central eminences. Such features are characteristic of young craters. It therefore appears that the tensile phase is recent; and the surface of the Moon may be in tension even at the present time. (This condition would seem to have a counterpart in the much discussed hypothesis of the expansion of the Earth's core, the hypothesis being based largely on the discovery of mid-oceanic ridges containing rifts which are thought to be tension features.)

The difficulty of the origin of the rilles, mentioned in section 1, is now automatically removed. It is well known that rilles are young features, and if the Ariadaeus and Hyginus Rilles owe their origin to wrench fractures[1], the rilles themselves would have formed only when the recent tensions in the Moon's crust opened the fractures, causing subsidence* of the surface.

4. THE HISTORY OF THE MOON'S CRUSTAL STRESSES

Recapitulating, it is thought that the oldest craters have been deformed so greatly from the

* Rilles themselves are emphatically not cracks, as many people seem to believe. See, *e.g.*, G. Fielder[7].

circular shape that the deviations from radial symmetry cannot be explained by tensions alone. It is suggested that the deformations arose as a consequence of thrust faulting. However, many recently-formed craters, at least, have originated during a surface tensile phase of the Moon, and the rilles—recent graben features—may also require crustal tensions for their formation.

All the observations may be explained by supposing that the crust of the Moon was in compression when the oldest visible craters were formed and that, more recently, the stresses reversed in sign, to become tensions. Such a reversal of sign could not occur suddenly. Hence, case (ii) of section 1 must be rejected, and it must be concluded that the *early compressive stress in the Moon's crust tended, generally, to decrease in magnitude as time elapsed.*

An estimate of the time required for this reversal of sign may be made, in principle, from two independent approaches. First, a suitably well-developed theory of erosion could be applied to estimate the time required to transform one of the young craters, discussed above, into one of the old. Secondly, attention might be paid to the possible time-scales of the compressive and tensile phases of the Moon's crust.

5. PROPOSED ORIGIN OF THE COMPRESSIVE AND TENSILE STRESSES

(i) An observed general *crustal compression* may be due to (a) the bulk expansion of the material of the crust relative to that of the core, or (b) the excess weight of a non-equilibrium bulge covering a large portion of the visible hemisphere (and, possibly, of the hemisphere which is averted from the Earth).

(ii) An observed general *crustal cracking* may only have its origin in an expansion of the material of the core relative to that of the crust.

A bulk expansion of any part of the Moon could result from radioactive heating. One can be reasonably sure that this process did operate, whereas other processes which have been listed[8] such as tidal forces and a possible (but not established) dependence of the "constant" of gravitation on time, are not known to have had any measurable effect on the Moon.

Case (a) could have operated as a result of radioactive heating only if the Moon had been differentiated as a whole, in order that the radioactive elements might be concentrated in a less dense crust. But, in this case, the core would have been at a higher temperature previously, and crustal cracking would have been *followed* by a crustal compressive phase. The reverse effects are observed. Consequently, with these assumptions, case (a) must be rejected. This means that the core of the Moon would be expected to be hotter than the crust, at the present time, and that the Moon has never been fully differentiated. It also leaves case (b) as the only mechanism capable of explaining the surface compressive phase. The neglect of processes other than radioactive heating means that these particular conclusions must be regarded merely as an hypothesis.

Warner[9] pointed out that the centre-line of any envelope bounding a fracture-pattern of a lunar crater (see also section 2) was always lined-up with one family of the grid system. This may now be readily understood; for, given that mechanism (i)b of section 5 produced compression failures which were approximately rectilinear when viewed in conical projection from the Earth, one would expect mechanism (ii), acting with similar uniformity of magnitude over the whole Moon, but in the opposite direction, to give rise to tension failures which followed the same general rectilinear pattern.

6. THE TIME SCALE

Two important facts, which arise directly from the observations, are that the actual deformation of a crater increased as time elapsed; and that the rate of deformation decreased as time elapsed. With the formation of craters, and the relief of crustal pressures by thrust-faulting, one might expect the smoothed pressure-time curve to fall exponentially.

Consider a time-scale on which the present time is $t=0$. Let a typical young crater have originated at time $t=-t_1$ and a typical old crater at $t=-t_2$, and let the present observed distortions of these craters be d_1, d_2 respectively (Fig. 5). Now $|t_2| > |t_1|$, and the curves for the old and young crater are, of necessity, iden-

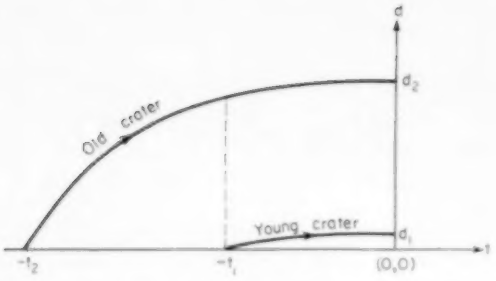


Fig. 5. The age of a crater and its distortion.

tical in form for $0 > t > -t_1$. It follows that

$$t_2 \gg \left(\frac{d_2}{d_1}\right) \cdot t_1. \qquad (3)$$

The values of the ratio d_2/d_1, obtained from the results given in Table 2, have been rounded off, in Table 4, to the nearest integer. A scale of relative times may be obtained by writing the inequality (3) as an equation. Then, if, one makes the assumption that the oldest craters originated $3 \cdot 10^9$ years ago, the ages of the other craters may be estimated. These assumptions were made in the construction of the age list given in Table 4, which does no more than suggest possible orders of magnitude of the ages of various groups of crater.

The conclusions of this paper, which are based purely on certain observations of the Moon's surface, appear to be in direct conflict with those reached by MacDonald[10] from a theoretical discussion of various models of a Moon heated by radioactivity. A task for the future will be to form a more rigorous theory which is compatible with observation.

Table 4.

Group	d_2/d_1	Suggested age (1 unit = 10^9 yrs.)	
		Old	Young
y_1	3	0·6	0·2
y_2	4	1·5	0·4
y_3	8	3·0	0·4

Acknowledgements—The writer wishes to thank T. Kiang, Th. Saemundsson, and B. Warner for useful discussions on these topics.

REFERENCES

1. G. Fielder, *J. Int. Lun. Soc.* **1,** 166 (1960).
2. K. von Bulow, *Geologie* (Berlin) **6,** 565 (1957).
3. K. von Bulow, *Beitr. Geophys.* **69,** 129 (1960).
4. G. P. Kuiper, (ed.) *Photographic Lunar Atlas*, Chicago University Press, Chicago (1960).
5. G. Fielder, *J. Brit. Astr. Ass.* **71,** in press (1961).
6. G. Fielder and B. Warner, *Mon. Nat. R. Astr. Soc.* to be published.
7. G. Fielder, *Structure of the Moon's Surface*, 11 and 12. Pergamon Press, London (1961).
8. G. Fielder, *Planet. Space Sci.*, **5,** 286, 1961
9. B. Warner, *J. Brit. Astr. Ass.* **71,** in press (1961).
10. G. J. F. MacDonald, *Planet. Space Sci.* **2,** 249 (1960).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 9–12. Printed in Great Britain.

THE NUMBER OF N_2 MOLECULES IN EARTH'S ATMOSPHERE FROM ROCKET DATA

S. N. GHOSH and K. D. SHARMA
J.K. Institute of Applied Physics, Allahabad University, Allahabad, India

(*Received* 28 *November* 1960)

Abstract—Using the data obtained from rocket-borne experiments, namely the intensities of solar X-rays in the different wavelength regions and depths of penetration of these radiations in the Earth's atmosphere, the amount of molecular nitrogen in the atmosphere was calculated. In the calculation it has been assumed that the energy received from the sun is spent wholly in ionizing O_2, O and N_2 molecules. The calculated number of N_2 molecules agrees with other available data.

1. INTRODUCTION

Although the distribution of O_2 molecules in the upper atmosphere is now well known, our knowledge of the distribution of N_2 molecules is meagre. The spectra of twilight airglow show the presence of N_2^+ ions from the first negative bands of N_2^+[1]. However, the number of such ions is very low. From rocket-borne experiments[2–7], the intensities of solar X-rays in the different wavelength regions and depths of penetration of these radiations in the Earth's atmosphere are known. Using these data, the amount of molecular nitrogen in the atmosphere above the depth of penetration is calculated, assuming that the energy received from the Sun is spent wholly in ionizing O_2, O and N_2 molecules. The results are then compared with other data[8, 9].

The intensities of solar X-rays in the different wavelength regions outside the Earth's atmosphere and depths of penetration of these radiations in the Earth's atmosphere as determined by rocket-borne experiments are collected in Table 1.

Table 1.

Wavelength region (Å)	Sun condition	Incident energy (erg/cm² sec)	Altitude h for 95% absorption (km)	Number of 0 atoms above altitude h (/cm² column)	Number of 0_2 molecules above altitude h (/cm² column)
6–10[2, 5]	160 minutes after class 1 flare	1×10^{-4} 1×10^{-3}	95	7×10^{18}	$1{\cdot}8\times10^{18}$
Below[7] 8	Late in Class 1 flare	5×10^{-3}	95	7×10^{18}	$1{\cdot}8\times10^{18}$
8–12[4]	Quiet	3×10^{-3}	100	$4{\cdot}9\times10^{18}$	$1{\cdot}5\times10^{17}$
8–18[5]	High coronal activity	0·6	105	$2{\cdot}6\times10^{18}$	$1{\cdot}8\times10^{16}$
8–20[4]	,,	0·1	105	$2{\cdot}6\times10^{18}$	$1{\cdot}8\times10^{16}$
8–20[6, 7]	Quiet	$1{\cdot}5\times10^{-3}$	105	$2{\cdot}6\times10^{18}$	$1{\cdot}8\times10^{16}$
8–20[7]	,,	$1{\cdot}3\times10^{-3}$	105	$2{\cdot}6\times10^{18}$	$1{\cdot}8\times10^{16}$
8–20[6, 7]	,,	$1{\cdot}2\times10^{-3}$	105	$2{\cdot}6\times10^{18}$	$1{\cdot}8\times10^{16}$
8–20[7]	Minimum solar activity	4×10^{-4}	105	$2{\cdot}6\times10^{18}$	$1{\cdot}8\times10^{16}$
44–60[6]	,,	$1{\cdot}4\times10^{-2}$	110	$2{\cdot}3\times10^{17}$	$5{\cdot}2\times10^{15}$
44–100[6]	,,	$2{\cdot}9\times10^{-2}$	120	5×10^{16}	$2{\cdot}8\times10^{13}$
44–100[6]	,,	$3{\cdot}5\times10^{-2}$	120	5×10^{16}	$2{\cdot}8\times10^{13}$

2. RATE OF PRODUCTION OF O^+ IONS

To illustrate the method of calculation of the number of N_2 molecules in the Earth's atmosphere, consider the data given in the first row of the above table.

Solar X-ray energy[2,5] in the region 6–10Å	5×10^{-4} erg/cm² sec
Mean energy of a photon in this region	$2{\cdot}64 \times 10^{-9}$ erg
Energy required for ejecting one *k*-electron from an O atom (critical absorption wave length is 23·58Å)	$8{\cdot}36 \times 10^{-10}$ erg
Energy of the ejected photo-electrons	$2{\cdot}64 \times 10^{-9} - 8{\cdot}36 \times 10^{-10}$ or $1{\cdot}8 \times 10^{-9}$ erg
Total number of O atoms[10] above 95 km (in reaching the altitude of 95 km about 95 per cent of these radiations is absorbed; the region for complete absorption is uncertain)	7×10^{15}/cm² column
Absorption coefficient[11] at 8·32Å	$4{\cdot}22 \times 10^{-20}$ cm²
Number of photons in the region 6–10Å	$5{\cdot}0 \times 10^{-4}/2{\cdot}64 \times 10^{-9}$ or $1{\cdot}9 \times 10^{5}$
Rate of formation of O^+ ions by *k*-shell ionization	$4{\cdot}22 \times 10^{-20} \times 7 \times 10^{15} \times 1{\cdot}9 \times 10^{5}$ or $5{\cdot}6 \times 10^{4}$/cm² column sec
Ionization potential of an O atom	$13{\cdot}61 \times 1{\cdot}6 \times 10^{-12}$ or $2{\cdot}18 \times 10^{-11}$ erg
Ions produced by one photoelectron	$1{\cdot}8 \times 10^{-9}/2{\cdot}18 \times 10^{-11}$ or 82·7 ions
Ions produced by $5{\cdot}6 \times 10^{4}$ photoelectrons	$5{\cdot}6 \times 10^{4} \times 82{\cdot}7$ or $4{\cdot}63 \times 10^{6}$ ions/cm² column sec
Total rate of O^+ ion production (due to *k*-shell ionization plus valency shell ionization)	$5{\cdot}6 \times 10^{4} + 4{\cdot}63 \times 10^{6}$ or $4{\cdot}7 \times 10^{6}$/cm² column sec
Energy required for this purpose	$(5{\cdot}6 \times 10^{4} \times 8{\cdot}36 \times 10^{-10}) + (4{\cdot}7 \times 10^{6} \times 2{\cdot}18 \times 10^{-11})$ or $1{\cdot}48 \times 10^{-4}$ erg.

3. RATE OF PRODUCTION OF O_2^+ IONS

Proceeding in the same way, the amount of energy utilized in the production of O_2^+ ions can be obtained. Remembering that the number of O_2 molecules above 95 km is $1{\cdot}8 \times 10^{18}$/cm² column[10]*, the absorption coefficient at 8·32 is $8{\cdot}44 \times 10^{-20}$ cm² and the ionization potential of O_2 is 12·2 eV, it can be shown that the total rate of O_2^+ ion production is $2{\cdot}7 \times 10^{6}$/cm² column sec and the energy utilized for this ionization is $7{\cdot}5 \times 10^{-5}$ erg.

4. RATE OF N_2^+ IONS PRODUCTION AND THE NUMBER OF N_2 MOLECULES ABOVE THE ALTITUDE OF PENETRATION

We have seen before that the energy used in the ionization of O and O_2 is $(1{\cdot}48 + 0{\cdot}75) \times 10^{-4}$ or $2{\cdot}23 \times 10^{-4}$ erg. Assume now that the surplus energy $(5 - 2{\cdot}23) \times 10^{-4}$ or $2{\cdot}77 \times 10^{-4}$ erg is utilized for the ionization of N_2. Now:

the energy required to eject one *k*-electron of N_2 (critical absorption wavelength for *k*-shell ionization of nitrogen is 31·18Å)[11]	$6{\cdot}38 \times 10^{-10}$ erg,

* Without much loss of accuracy, it may be assumed that the absorption coefficients of O_2 and N_2 are twice the absorption coefficients of O and N atoms respectively.

the energy of the ejected photo-electron	$(2{\cdot}64\times10^{-9}-6{\cdot}36\times10^{-10})$ or 2×10^{-9} erg,
the ionization energy of N_2	15·58 eV or $2{\cdot}495\times10^{-11}$ erg,
one photo-electron produces	$2\times10^{-9}/2{\cdot}5\times10^{-11}$ or 80·5 ions,
if x be the rate of k-electron production, the total energy used in the ionization of N_2 is	$x\,(6{\cdot}36\times10^{-10}+2\times10^{-9})=2{\cdot}77\times10^{-4}$ or, $x=1{\cdot}05\times10^{5}/\text{cm}^2$ column sec,
the total rate of N_2^+ ions production is	$1{\cdot}05\times10^{5}+80{\cdot}5\times1{\cdot}05\times10^{5}$ or, $8{\cdot}45\times10^{6}/\text{cm}^2$ column sec,
hence the number of N_2 molecules above 95 km is	$\dfrac{1{\cdot}05\times10^{5}}{1{\cdot}9\times10^{5}\times5{\cdot}16\times10^{-20}}$ or $2{\cdot}14\times10^{19}$ atoms/cm^2 column or $1{\cdot}07\times10^{19}$ molecules/cm^2 column.

Following the method illustrated above and using the data given in Table 1 and the absorption coefficients in Table 2, the number of N_2 molecules above different altitudes are calculated. The results are given in Table 3 (ionization due to Auger Effect is very small and can be neglected[12]). Comparing with other available data (Table 4), it is found that the agreement is quite good.

Table 2. Atomic absorption coefficients at different wavelengths in the X-ray region

Wavelength (Å)	Atomic absorption coefficient of O (cm^2)	Atomic absorption coefficient of N (cm^2)
1·537	$3{\cdot}3\times10^{-22}$	$1{\cdot}7\times10^{-22}$
1·935	$5{\cdot}9\times10^{-22}$	—
2·287	$9{\cdot}82\times10^{-22}$	—
3·378	$3{\cdot}11\times10^{-21}$	$1{\cdot}85\times10^{-21}$
3·592	$3{\cdot}75\times10^{-21}$	$2{\cdot}22\times10^{-21}$
3·927	$5{\cdot}05\times10^{-21}$	$2{\cdot}78\times10^{-21}$
4·146	$5{\cdot}9\times10^{-21}$	$3{\cdot}5\times10^{-21}$
4·359	$6{\cdot}85\times10^{-21}$	$3{\cdot}86\times10^{-21}$
5·166	$1{\cdot}09\times10^{-20}$	$6{\cdot}4\times10^{-21}$
5·394	$1{\cdot}26\times10^{-20}$	$7{\cdot}3\times10^{-21}$
6·973	$2{\cdot}59\times10^{-20}$	$1{\cdot}5\times10^{-20}$
8·323	$4{\cdot}22\times10^{-20}$	$2{\cdot}58\times10^{-20}$
9·868	$6{\cdot}75\times10^{-20}$	$4{\cdot}17\times10^{-20}$
23·58	$6{\cdot}48\times10^{-19}$	—
31·18	—	$7{\cdot}28\times10^{-19}$
43·5	$1{\cdot}54\times10^{-19}$	$8{\cdot}8\times10^{-18}$
68	$4{\cdot}32\times10^{-19}$	$2{\cdot}33\times10^{-19}$

Table 3.

Rate of O^+ ion production above altitude h (/cm² col. sec)	Rate of O_2^+ ion production above altitude h (/cm² col. sec)	Rate of N_2^+ ion production above altitude h (/cm² col. sec)	Number of calculated N_2 molecules above altitude h (/cm² column)
$5{\cdot}3\times10^{6}$	$2{\cdot}7\times10^{6}$	$8{\cdot}5\times10^{6}$	$1{\cdot}1\times10^{19}$
$7{\cdot}3\times10^{6}$	$4{\cdot}4\times10^{6}$	$1{\cdot}6\times10^{8}$	$1{\cdot}5\times10^{20}$
$2{\cdot}8\times10^{7}$	$1{\cdot}9\times10^{6}$	$5{\cdot}3\times10^{7}$	$7{\cdot}7\times10^{18}$
$5{\cdot}7\times10^{9}$	$8{\cdot}7\times10^{7}$	$9{\cdot}7\times10^{9}$	$4{\cdot}3\times10^{18}$
$8{\cdot}9\times10^{8}$	$1{\cdot}4\times10^{7}$	$1{\cdot}6\times10^{9}$	$4{\cdot}3\times10^{18}$
$1{\cdot}1\times10^{7}$	$1{\cdot}7\times10^{5}$	$2{\cdot}3\times10^{7}$	$4{\cdot}4\times10^{18}$
$9{\cdot}4\times10^{6}$	$1{\cdot}5\times10^{5}$	2×10^{7}	$4{\cdot}3\times10^{18}$
$8{\cdot}7\times10^{6}$	$1{\cdot}3\times10^{5}$	$1{\cdot}8\times10^{7}$	$4{\cdot}3\times10^{18}$
$2{\cdot}9\times10^{6}$	$4{\cdot}5\times10^{4}$	6×10^{6}	$4{\cdot}3\times10^{18}$
4×10^{7}	$2{\cdot}1\times10^{5}$	$5{\cdot}2\times10^{8}$	$3{\cdot}4\times10^{18}$
$3{\cdot}5\times10^{7}$	$4{\cdot}6\times10^{4}$	$1{\cdot}1\times10^{9}$	$1{\cdot}9\times10^{18}$
4×10^{7}	$5{\cdot}4\times10^{4}$	$1{\cdot}4\times10^{9}$	$1{\cdot}9\times10^{18}$

Table 4. Number of molecular nitrogen in the atmosphere above the depth of penetration.

Altitude h (km)	Number of N_2 molecules above altitude h[8] (/cm² column)	Number of N_2 molecules above altitude h[9] (/cm² column)	Calculated number of N_2 molecules above h (/cm² column)
95	$4{\cdot}0\times10^{19}$	$3{\cdot}8\times10^{19}$	$1{\cdot}1\times10^{19}$
100	$1{\cdot}4\times10^{19}$	$1{\cdot}2\times10^{19}$	$7{\cdot}7\times10^{18}$
105	$8{\cdot}3\times10^{18}$	$7{\cdot}0\times10^{18}$	$4{\cdot}3\times10^{18}$
110	$3{\cdot}2\times10^{18}$	$2{\cdot}6\times10^{18}$	$3{\cdot}4\times10^{18}$
120	$9{\cdot}0\times10^{17}$	$7{\cdot}0\times10^{17}$	$1{\cdot}9\times10^{18}$

REFERENCES

1. M. Dufay, *Ann. Phys. Paris* **8**, 813 (1953).
2. H. Friedman, S. W. Lichman and E. T. Byram, *Phys. Rev.* **83**, 1025 (1951).
3. R. Tousey, K. Watanabe and J. D. Purcell, *Phys. Rev.* **83**, 792-797 (1951).
4. T. R. Burnight, *Physics and Medicine of the Upper Atmosphere*, p. 226, University of Mexico Press, Albuquerque (1952).
5. E. T. Byram, T. A. Chubb and H. Friedman, *Rocket Exploration of the Upper Atmosphere*, p. 274 Pergamon Press, London (1954).
6. E. T. Byram, T. A. Chubb and H. Friedman, *J. Geophys. Res.* **61**, 251 (1956).
7. T. A. Chubb, H. Friedman, R. W. Kreplin and J. E. Kupperian Jr., *J. Geophys. Res.* **62**, 389 (1957)
8. H. K. Kallman, W. B. White and H. E. Newell, *J. Geophys. Res.* **61**, 513 (1956).
9. J. E. Miller, *J. Geophys. Res.* **62**, 351 (1957).
10. D. R. Bates, Witherpoon and E. Agnes, *Mon. Nat. R. Astr. Soc.* **112**, 101 (1952).
11. A. Compton and S. K. Allison, *X-rays in theory and Experiment*, Table No. VII-8, p. 524 and Table No. I of Appendix.
12. S. N. Ghosh and S. Nand. *Indian J. Phys.* **34**, 516 (1960).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 13–22. Printed in Great Britain.

LONGITUDINAL DEPENDENCE OF RADIATION–BELT SCATTERING, AND PRIMARY AURORAL PARTICLES

C. J. LOUGHNAN
Dominion Physical Laboratory Auroral Station, Omakau, New Zealand

(*Received* 4 *April* 1961)

Abstract—Due to the geomagnetic field differing from that of a centred dipole, the altitude of the trapped-radiation regions surrounding the earth varies with geomagnetic longitude. In some circumstances this should lead to a marked longitude dependence in the leakage of particles from these regions. This dependence is estimated, and also the precipitation pattern for the outer radiation belt. The conditions for the existence of this variation are discussed, and a consequent difficulty is raised in explaining aurorae as the direct result of leakage from the radiation belt. As a preliminary, a simple method of estimating the altitude of surfaces of constant magnetic field strength is presented.

1. INTRODUCTION

The earth is surrounded by an extensive corona of fast electrons and protons trapped within its magnetic field and forming the Van Allen radiation[1]. Particular attention will here be given to the outer radiation belt, which consists of 10 to 100 keV electrons trapped on the lines of force that leave the earth within 5 or 10° of 60° geomagnetic latitude. The lowest regions of this belt reach down to the order of 1,000 km altitude, and electrons leaking from these regions are considered by some to be primary auroral particles. The motion of a trapped particle consists of: a very rapid gyration around a magnetic field line; a fast oscillation or bounce from hemisphere to hemisphere along a field line, bounded by magnetic reflection at a particular field strength; and a slow (longitudinal) drift around the earth[2]. In a static magnetic field, with no electric field, the only other motion is scattering by the ambient atmosphere, almost all of which occurs at the particle's lowest altitude. Scattering, and presumably other processes associated with magnetic disturbance[3], determines the trapped lifetime.

Because of the asymmetry of the geomagnetic field, the radiation belts penetrate to different altitudes at different longitudes. Thus the atmospheric scattering experienced by particles near their mirror points is a function of longitude. In the next section a simple method of estimating the altitude of constant-magnetic-field-strength contours is examined; this is needed to locate particle reflection points. Then the longitude dependence of atmospheric scattering is derived, followed in section 4 by the pattern of outer-belt precipitation. In these sections it is assumed that the drift in longitude is sufficiently rapid to prevent any local depletion of the radiation belt at the regions of greatest scattering: in section 5 this assumption is examined by comparing drift periods and particle lifetimes. Finally the possibility that scattered particles explain aurorae is discussed with reference to the longitude effect.

2. ALTITUDE OF CONSTANT-FIELD-STRENGTH SURFACES

The reflection points of a particle's orbit fall on two closed curves which would be parallels of latitude were the geomagnetic field that of a dipole. These reflection-point curves, or "drift paths," projected onto the earth's surface, are available from Vestine and Sibley's paper[4]. We require also the altitude variation of these drift paths, which are at a constant magnetic field strength. Thus a method is required for estimating the altitude to a particular field strength, a method applicable for altitudes of the order of 1000 km.

The altitude to a particular field strength may be determined from the values of the field and its derivatives at the earth's surface. It will be written down in an exact form, and then various simple approximations will be considered.

Using a Taylor-series expansion, the altitude h to a total field strength F above a specified place is

$$h = -(F_0 - F)(dr/dF)_0 + \tfrac{1}{2}(F_0 - F)^2 (d^2r/dF^2)_0 - \quad (1)$$

where r is the radial distance from the earth's centre, and the zero suffix denotes values at the earth's surface.

In a centred-dipole field the differentials are

$$(dr/dF)_0 = -r_0/3F_0, (d^2r/dF^2)_0 = 4r_0/9F_0^2. \quad (2)$$

Thus a first approximation to h is

$$h_1 = r_0 (F_0 - F)/3F_0 \quad (3)$$

where the first term of equation (1) is used with the dipole value for the gradient. An improved value for h is obtained by taking the second order term while retaining dipole values for the differentials:

$$h_2 = h_1 + 2r_0 (F_0 - F)^2/9F_0^2 \quad (4)$$

An alternative improvement is to take the first order term with the correct, non-dipole, differential:

$$h_3 = -(F_0 - F)(dr/dF)_0 \quad (5)$$

where the differential can be shown to be given by

$$-(r_0/F_0)(dF/dr)_0 = 2 - \cos^2 I + \delta I/\delta\theta - \sin I \cos I (\delta D/\delta\phi) + \tan\lambda \sin I \cos I \cos D \quad (6)$$

I is the dip angle, D the declination angle, λ the true latitude, and θ and ϕ are co-ordinates, along magnetic north and west respectively, measured as geocentric angles. Equations (5) and (6) give the first order approximation to h from any part of the surface with no influence from other regions, and so leave this method free from the effects of smoothing involved in using harmonic analysis.

The above three estimates of h may be obtained readily from the magnetic field at the earth's surface, the first two requiring only the values of F_0 there. These heights have been calculated for the 0·5 gauss surface, at intervals of 15° of *geographic* longitude along Vestine and Sibley's auroral zone[4], and the results are shown in Fig. 1. We see that the crudest approximation, h_1, gives the general behaviour of h quite well. The addition of the dipole second-order term is seen to be of greater importance than the use of the exact first-order term. A (dipole) third-order term makes hardly any further difference. For comparison,

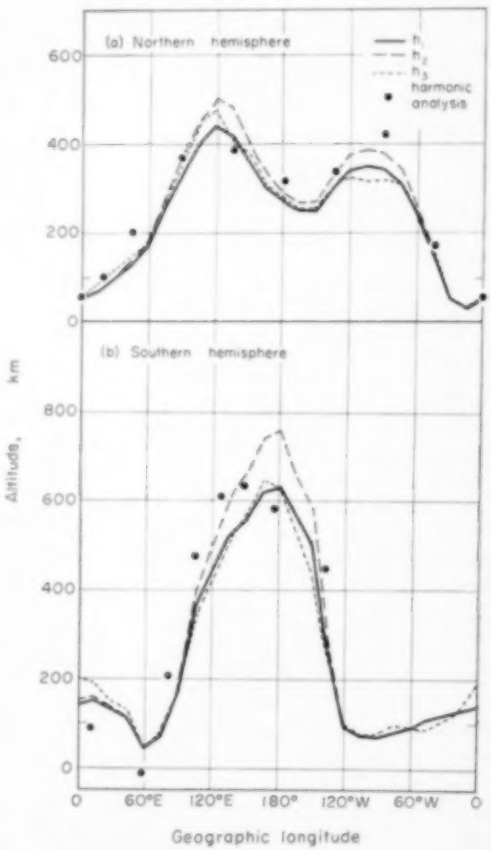


Fig. 1. Height of 0·5 gauss surface above Vestine and Sibley's theoretical auroral zone [4], from the field strength at the earth's surface. h_1, first-order dipole approximation. h_2, second-order dipole approximation. h_3, exact first-order term. Also heights from a harmonic analysis[4].

Fig. 1 also shows heights determined from a 48-term spherical-harmonic analysis[4].

For the next section of this paper the first approximation, h_1, is adequate and of course it is extremely easy to determine. However a variant of this approximation was used. For the altitude of the surface $F=\overline{F}_0$, where the field strength equals the average value on the earth's surface beneath the appropriate drift-path, this gives

$$h=r_0(F_0-\overline{F}_0)/3\overline{F}_0.$$

This equation gives a mean height of zero. It is assumed that at about 1000 km altitude the variation of h is much the same as at the surface. Thus the variation Δh of the altitude from its mean value is approximated by

$$\Delta h=r_0(F_0-\overline{F}_0)/3\overline{F}_0 \qquad (7)$$

At much greater altitudes the height variation tends to that given by the quadrupole terms of the earth's field, the major portion of which is included in the eccentric dipole approximation to the field. This last variation is of similar order to that of Fig. 1, but distributed around the globe in a very different manner.

3. CALCULATION OF THE LONGITUDE EFFECT

The reflection points of a particle drift around a pair of closed curves or "drift paths," one in each hemisphere. Calculations of these paths are available[4]. Along any pair of paths the variation of reflection-point altitude is readily obtained using equation (7). It is assumed that along any reflection-point locus the trapped-particle density is constant, and that the scattering down is simply proportional to the atmospheric density encountered. For a reflection-point altitude h, and an atmospheric scale height H, the atmospheric density and precipitated flux are given by a "longitude factor" $\exp\{-(h-h_0)/H\}$, where h_0 is some datum level to be fixed. For H a value of 170 km, independent of altitude and latitude, is chosen. The correct value is not well enough known to contradict this figure, which may be obtained from a paper by Nicolet[5] for 1000 km altitude by halving his value of 0·6 ergs cm^{-2} for vertical heat-conduction so as to get agreement with satellite data[6]. This value of H assumes that the atmosphere is essentially atomic oxygen and nitrogen; where hydrogen predominates H will be about 16 times greater and the exponential factor will vary only slightly around the globe. Here it is convenient to choose the reference altitude h_0 to be the mean value of h. Thus $\exp(-\Delta h/H)$ provides the desired longitude factor, Δh being taken from equation (7). The factor is unity at the average height of a drift-path pair; and it gives the global flux distribution from a radiation belt whose bottom surface has the same average height over every such pair of drift paths. Such a lower boundary is very closely the same as that of a radiation belt which, when averaged over all geomagnetic longitudes, has a lower border independent of geomagnetic latitude.

This longitude factor, $\exp(-\Delta h/H)$ where Δh comes from equation (7) and H is 170 km, is plotted in Fig. 2. One may note the extremely large values over the magnetic anomaly in the south Atlantic.

4. THE PRECIPITATION PATTERN FROM THE OUTER RADIATION BELT

It is a simple matter to extend the method of the last section to obtain the distribution of particles scattered from the radiation belts. The result of doing this for the outer belt is shown in Fig. 3. The belt's lower boundary is given as a function of geomagnetic latitude in Fig. 8 of reference[7]. Since in Fig. 2 pairs of drift paths were used, and not circles of geomagnetic latitude, the average lower boundary for a given geomagnetic latitude is assumed the same as that for the drift-path pair of the same average geomagnetic latitude. The precipitation pattern is obtained by multiplying the longitude factor of the last section by a latitude-dependent factor $\exp(-d/H)$. Here H is the scale height (170 km), and d is the average altitude of the lower surface of the radiation belt as a function of geomagnetic latitude. The zero level is chosen to make d zero where the belt, after being averaged in longitude, penetrates deepest into the atmosphere.

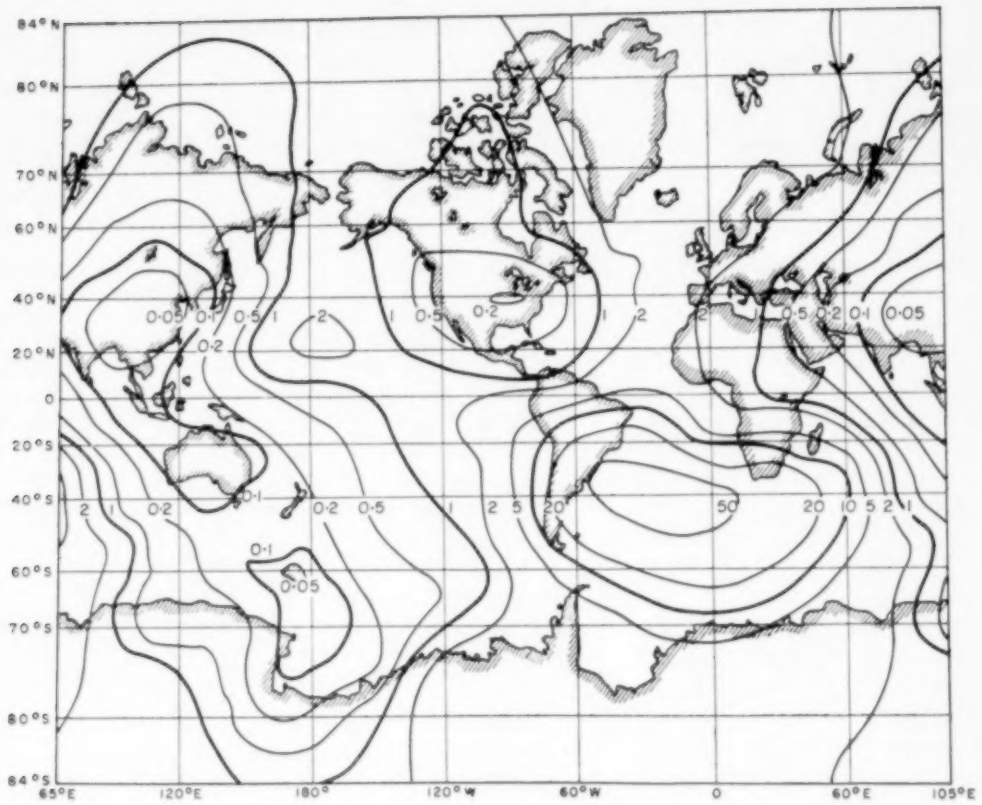


Fig. 2. The longitude factor.

5. PARTICLE ENERGIES REQUIRED TO PRODUCE THE LONGITUDE EFFECT

For scattering to have the longitudinal asymmetry estimated above, trapped particles must be redistributed in longitude rapidly enough to prevent an excess depletion of the radiation belts over the regions of greatest scattering. If the particle lifetimes are very much greater than the period of longitudinal drift, the density of their reflection points must be uniform along the drift paths. In this case the scattered flux must vary as the atmospheric density, and be distributed in the asymmetric manner estimated in the previous sections. If, on the other hand, particles drift only a negligible amount during their lifetime, their lives will be relatively short over regions of low field strength, and therefore their density (on a given constant-field-strength surface) will be relatively small over such regions. In this case there is no longitude effect. To obtain such a uniform precipitation, the lowest intensity-contours of the radiation belt must be at constant altitude instead of constant field strength. A simple criterion is to require particles to drift once or more around the globe while trapped, if they are to produce the longitudinal asymmetry.

In this section longitudinal drift periods and particle lifetimes are estimated. The analysis of Alfvén[8] is followed, except that, unless otherwise stated, the rationalized M.K.S. system of units is used.

The symbols used are:

r radial co-ordinate, from earth's centre.
λ geomagnetic latitude.

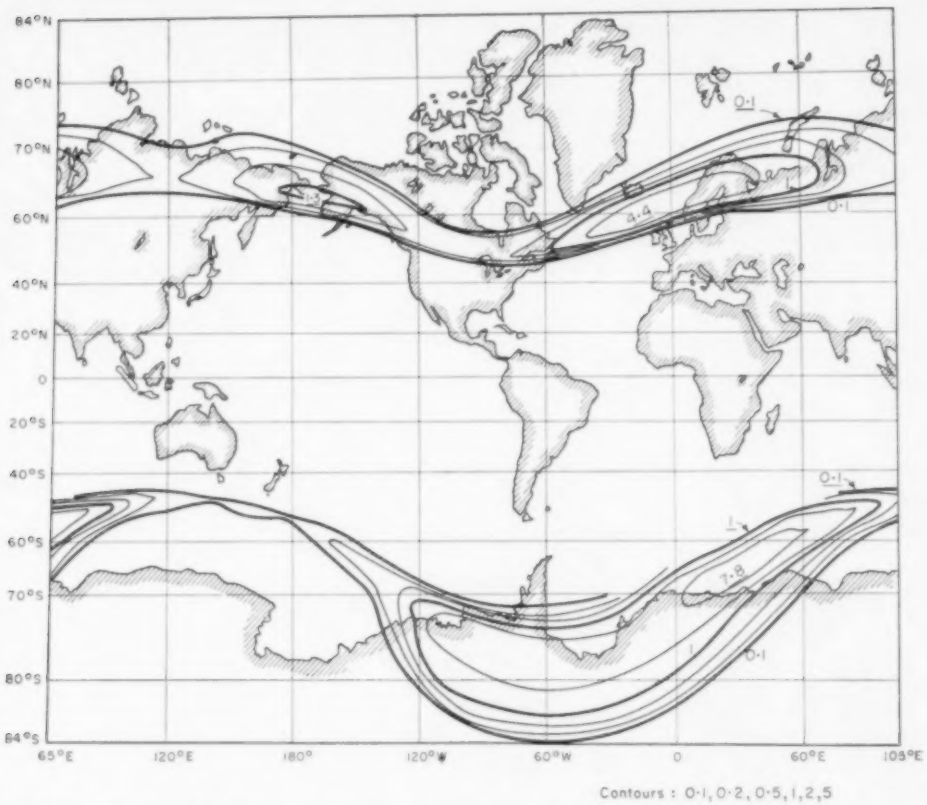


Fig. 3. The precipitation zones for the outer radiation belt, showing the longitude effect.

$\mathbf{r}_e$ equatorial radius vector to an orbit.
m The suffix m distinguishes reflection-point co-ordinates.
t_b Period of interhemispheric oscillation or bounce.
t Period of longitudinal drift.
T_m Lifetime a trapped particle would have if it were always at its reflection point.
T Lifetime of a trapped particle.
η $\cos^{-6}\lambda\sqrt{1+3\sin^2\lambda}$.
I, J, K Functions of λ_m, defined as:

$$I = 3\int_0^{\lambda_m} \frac{\cos^3\lambda\,(1+\sin^2\lambda)\,(1-\eta/2\eta_m)}{(1+3\sin^2\lambda)^{3/2}\,(1-\eta/\eta_m)}\,d\lambda,$$

$$J = \int_0^{\lambda_m} \frac{\cos\lambda\,(1+3\sin^2\lambda)^{1/2}}{(1-\eta/\eta_m)^{1/2}}\,d\lambda,$$

$$K = \pi I^{-1} J \cos^2\lambda_m.$$

H Atmospheric scale height.
ρ Atmospheric mass density.
N Atmospheric particle density in cm^{-3}.
Z Nuclear charge of an atmospheric atom.
e Electronic charge ($1{\cdot}6\times10^{-19}$ coulombs).
a Geomagnetic dipole moment ($8{\cdot}1\times10^{15}$ Weber metres).
E, v Energy and speed of a fast particle.
E_0 Rest mass of electron in energy units.
σ Cross-section.
C_{St} Störmer's unit of length (reference (8), section 2.4).
Θ Scattering angle.

To discuss drift periods and lifetimes the following simplifications are assumed: the geomagnetic field is that of a centred dipole; the

important atmospheric regions consist of atomic oxygen and atomic hydrogen, H being 170 km in the former and 16 times as great, viz. 2700 km, in the latter. The transition from oxygen to hydrogen may be placed at about 1200 km[9]. Only sub-relativistic particles are considered.

Because a dipole field is used, on any orbit

$$r_m = r_e \cos^2 \lambda_m.$$

The bounce and drift periods may be written down from section 2.5 of Alfvén[8]. Rewriting his equation (15):

$$t_b = (2\, ear_e / EC_{St}^2)\, J \qquad (8)$$

Equation (18) gives the number of bounce cycles in a drift period

$$(t/t_b) = \tfrac{1}{2}\pi\, (C_{St}/r_e)^2\, I^{-1} \qquad (9)$$

Multiplying our equations (8) and (9) together:

$$t = (\pi ea/Er_e)\,(J/I)$$
$$= (ea/Er_m)\, K \qquad (10)$$

The function K, giving the dependence of drift period on latitude, is plotted in Fig. 4. It is obtained taking I from Alfvén's Fig. 2.4[8], and calculating J—the value at 60° latitude being found by graphical integration. $K \sec^2 \lambda_m$ varies only from 2·1 at the equator to π at the poles, so K varies similarly to $\cos^2 \lambda_m$.

Inserting numerical values in equation (10), with r_m taken as 7500 km for the lower regions of the radiation belt, there results

$$t = 13\ K/E \text{ days } (E \text{ in keV}),$$

and at 60° latitude this is

$$t = 10E^{-1} \text{ days } (E \text{ in keV}) \qquad (11).$$

This last is plotted as the line I on Fig. 5.

Lifetimes from scattering. The time spent by a particle in a trapped orbit is desired, for comparison with its drift period. This lifetime, T, will now be determined, assuming that the only loss mechanism is interaction with the ambient atmosphere. The time constant T_m for the rate of loss while the particle is at its reflection point

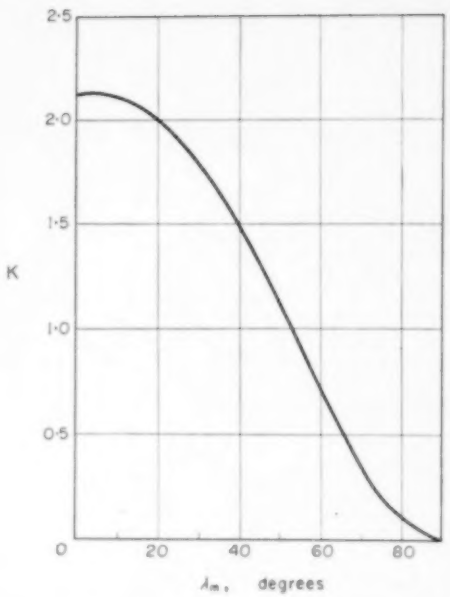


Fig. 4. The function K, giving the dependence of drift period on reflection-point latitude.

is more readily obtained, and from it T is obtained since

$$T = (\rho_m/\bar{\rho})\, T_m \qquad (12)$$

Here $\bar{\rho}$ is the average atmospheric density along the orbit. The function $\rho_m \sqrt{r_m H}/\bar{\rho}$, a function of λ_m alone, has been calculated by Ray[10].

Firstly T_m will be determined for the trapped electrons. The method is somewhat similar to that of Welch and Whittaker[11]. The loss of electrons is by a gradual Coulomb scattering, most of which occurs very near the reflection points where the greatest atmospheric density is encountered. Since at these points the orbit is normal to the magnetic field, any scattering alters the pitch angle and lowers the reflection point. The desired T_m may be defined as the time required for this scattering, for a particle always at its reflection point, to lower the reflection point by a scale height. Let the pitch angle of an electron's spiral orbit, originally 90°, change by Θ due to the scattering. This change lowers the reflection point a distance H if

$$\Theta = \pm (3H/r_m)^{1/2}$$

the magnetic field being taken as vertical. Thus T_m may be defined as the time required for scattering such that

$$\overline{\Theta^2} = 3H/r_m \qquad (13)$$

Substituting this value of $\overline{\Theta^2}$ into equation (2.29) of reference (12), and substituting all possible numerical values (including $r_m = 7500$ km), one obtains:

$$T_m = \frac{0{\cdot}30\, HE^{3/2}(E+2E_0)^{3/2}}{NZ^2(E+E_0)\ln\left\{\dfrac{18E(E+2E_0)}{Z^{4/3}(E+E_0)}\right\}} \text{ days} \qquad (14)$$

where E is in keV ($E_0 = 510$), H in km and N in cm^{-3}.

In an oxygen atmosphere Z is 8 and H 170 km, so

$$T_m = \frac{0{\cdot}78\, E^{3/2}(E+2E_0)^{3/2}}{N(E+E_0)\ln\{1{\cdot}15E(E+2E_0)/(E+E_0)\}} \text{ days} \qquad (15)$$

In hydrogen Z is unity and H 2700 km, giving

$$T_m = \frac{830\, E^{3/2}(E+2E_0)^{3/2}}{N(E+E_0)\ln\{18E(E+2E_0)/(E+E_0)\}} \text{ days} \qquad (16)$$

Finally, the lifetime T may be found from T_m using equation (12), and the published values[10] of $\rho_m\sqrt{r_m H}/\bar{\rho}$.

This leads to the graphs of T for electrons, plotted in Fig. 5 as curves II and III for atmospheres of oxygen and hydrogen respectively. These curves are actually graphs of $10^{-4}\,NT$ for a latitude of 60°, and give the lifetimes in an atmospheric density of 10^4 atoms cm^{-3} directly.

For trapped protons the time T_m is determined by charge exchange with the ambient atmosphere. This results in a fast hydrogen atom and a slow ion, so that the proton, unlike an electron, is suddenly lost to the radiation belt. If σ is the cross-section for the charge-exchange reaction, then, in any consistent system of units,

$$T_m = (N\sigma v)^{-1}. \qquad (17)$$

The lifetime T may readily be found, now that T_m is defined, using equation (12) as was done for electrons.

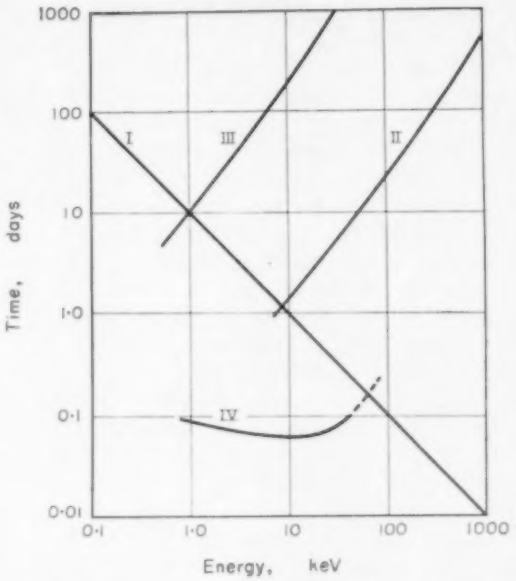


Fig. 5. Drift periods, and trapped lifetimes from atmospheric scattering.
I τ, the period of longitudinal drift. II, III, IV The lifetimes T of, respectively, trapped electrons in an oxygen atmosphere, electrons in hydrogen, and protons in hydrogen, for an atmospheric density of 10^4 cm^{-3}. For a density of N cm^{-3}, the graphs give 10^{-4}NT.

In the case of an atmosphere of atomic hydrogen there is information available concerning σ. This is discussed by Fite *et al.*[13]; their experimental values lead to curve IV of Fig. 5, which presents $10^{-4}\,NT$ for trapped protons reflected at latitude 60° in an atmosphere of N hydrogen atoms per cubic centimetre. For an oxygen atmosphere values of σ do not seem to be available. Should these values be about the same as for hydrogen, which is not unreasonable, curve IV of Fig. 5 would apply for either atmospheric constituent.

Drift required to produce the longitude effect. The longitudinal dependence calculated in Sections III and IV will exist only if the density of reflection points is independent of longitude. To keep the reflection points uniformly distributed, in spite of non-uniform scattering, an adequate drift in longitude is required. A simple criterion is to demand that a particle drifts at least once around the globe in its life-

time, i.e. that T exceeds t, if the longitude effect is to be produced.

The scale height of atmospheric hydrogen is much greater than the height variation of constant-field-strength surfaces, so that scattering by hydrogen cannot exhibit more than a slight dependence on longitude.

The lifetime T, that is to be compared with t, must therefore be the lifetime in the oxygen region of the atmosphere, and now a value for the atmospheric density must be inserted in the equation for T. The composition and density of the atmosphere are here taken from Fig. 3 of reference (9), and the heavy component is taken to be oxygen atoms even though in fact there may be nitrogen also. At about 1200 km altitude the atmosphere changes from largely oxygen atoms to largely hydrogen atoms, and in the vicinity of this altitude there are about 10^4 hydrogen atoms per cubic centimetre. For a given atmospheric density the life of trapped protons may be much the same in oxygen as in hydrogen, in which case protons are mainly scattered by oxygen below about 1200 km. Trapped electrons mirroring in about 10^2 cm^{-3} of oxygen have a lifetime equal to their life in 10^4 cm^{-3} of hydrogen (Fig. 5), and the densities of oxygen and hydrogen have these values, respectively, at an altitude slightly above 2000 km (Actually at this height the oxygen is present as ions rather than neutral atoms). Below this height trapped electrons are mainly scattered by oxygen rather than by hydrogen.

Assuming for the moment that particles are first trapped into orbits that do not reach down to the oxygen atmosphere, the lifetime T may be estimated readily. Protons may be expected to disappear after a collision without first diffusing down into the oxygen; thus they are not expected to exhibit a longitude effect. To produce the effect protons must not only mirror below 1200 km, but as can be seen from Fig. 5 they must have energies greater than some 60 keV if their reflection points are near 60° latitude. Electrons will be scattered and their reflection points will eventually enter the oxygen regions at slightly above 2000 km altitude. Entry of the reflection points into the oxygen region may be defined as occurring when the lifetimes against oxygen and hydrogen scattering are equal, and this gives the remaining lifetime, the desired value of T, as half the lifetime determined from oxygen or hydrogen scattering alone. That is T is one half of the lifetime of an electron being reflected in about 10^4 cm^{-3} of hydrogen, or one half of the values given by curve III in Fig. 5. Thus, if T is determined by scattering and the initial points are high enough, the value of T may now be written down. Inspection of Fig. 5 shows that at latitude 60°, T exceeds t for electrons of over 1 keV energy. This critical energy decreases somewhat towards high latitudes. Thus scattered electrons of over 1 keV should show the longitude effect, while electrons of lower energy should come down uniformly. The density of trapped electrons of less than 1 keV penetrating to about or below 2000 km should vary with longitude; in fact instead of being uniformly distributed along a drift path they should be uniformly distributed along the projection of a drift path onto a surface of constant altitude. The energy range above 1 keV includes the particles observed in the radiation belts, so they are expected to be scattered in the manner derived in sections III and IV. Electrons from the outer belt should therefore be scattered over the zones shown in Fig. 3. This agrees with the observed radiation-belt contours being symmetric with respect to the geomagnetic field[7], not being constant-altitude contours at the lowest levels as would be expected if the longitudinal drift was insufficient.

Above it was assumed that electrons are trapped above the region where scattering by oxygen is important. It should be pointed out that the average reflection height is a few hundred kilometres above the effective height of scattering. The height at which the atmospheric density equals the effective value N (which is the average density along the drift paths), while above the minimum reflection height, can be a few hundred kilometres below the average reflection height. For electrons to be trapped above the region of oxygen scattering their initial reflection points must exceed some 2500 km in average altitude. Alternatively, they must be injected above 3000 to 3500 km, because if one assumes an isotropic distribution at injection

the average injection and reflection altitudes can readily be shown to differ by about 700 km. Injection at low altitudes is possible from cosmic ray albedo neutrons[14], but even from this source most electrons are injected well above the atmospheric oxygen.

Other dumping processes. Professor J. A. van Allen (private communication) has pointed out that particles may be lost from the radiation belt more rapidly than can be explained by atmospheric scattering. There are observations of sudden dumping of electrons during the early stages of magnetic storms[3, 15]. If 20 keV is assumed to be the typical energy of outer-belt electrons, Fig. 5 shows that the drift period is half a day. Dumping processes which cause reflection points to traverse the regions of scattering by oxygen in half a day or less must therefore be expected to lead to precipitation that is independent of longitude, and the dumping early in a magnetic storm seems quite rapid enough to do this.

To sum up, the following conclusions were reached in this section. Protons reflected above 1200 km are lost without any longitudinal asymmetry, since any longitude effect produced by scattering by hydrogen may be neglected. For any protons being reflected below 1200 km and for electrons there is a critical energy to be considered. Above this energy particles drift with sufficient rapidity to exhibit the longitude effect, being dumped in the manner described by Figs. 2 and 3, and (as observed[7]) the density of reflection points is determined by the magnetic field independently of longitude. Below the critical energy particle precipitation should be independent of longitude, and the lower contours of the reflection-point density should be at constant altitude rather than constant field strength. For particle loss due essentially to atmospheric interactions, and at a latitude of 60° geomagnetic, the critical energies have been estimated as 60 keV for protons and 1 keV for electrons. As the outer belt contains electrons of well over 1 keV, a longitude effect is expected in this case. However, on occasions a rapid dumping of outer belt electrons has been observed. When this happens the loss is fast enough to raise the critical energy above that typical of outer-belt electrons, and no longitude effect is then to be expected.

6. THE LONGITUDE EFFECT AND PRIMARY AURORAL PARTICLES

It is important to find out how the leakage of particles may become apparent where they are absorbed in the lower regions of the ionosphere. It has been proposed by van Allen, McIlwain and Ludwig[7] that visible aurorae are the direct result of the precipitation of trapped particles, which it seems may now be particularized as electrons from the outer belt. However, it is difficult to reconcile this with what has already been said about the dependence of precipitation on longitude. The overall occurrence of aurorae is not distributed as in Fig. 3, the auroral isochasms actually approximating circles of geomagnetic latitude or, better[4], trapped-particle drift paths. Only the regions from the auroral zones equatorward are being considered here, since only they are beneath the trapping region. As will be recalled from the last section, electron loss from the radiation belt must be due to rapid dumping processes if a suitably symmetrical precipitation pattern is to be obtained. Aurorae occur very frequently under the higher-latitude regions of the belt, being observed every night at the high-latitude boundary. On present information it seems that sudden dumping cannot occur often enough to permit the explanation of aurorae by radiation-belt precipitation. Only relatively low latitude auroral displays can be so explained without implying a longitudinal asymmetry.

Thus the hypothesis that a flux of primary electrons from the radiation belt is the direct cause of auroral emission must be considered inadequate, on the evidence of the simple and gross feature of distribution in longitude. Apart from this, other sources of auroral excitation cannot be excluded while it is unknown whether or not the radiation-belt leakage can produce the essential features of the aurora and in particular the fine scale of its spatial structure. It remains, of course, to propose the interesting possibility that some auroral phenomena, unlike the total occurrence of all types of visible display, are distributed as in

Fig. 3. It is to be hoped that I.G.Y. synoptic data will lead to the necessary information here.

Another possible example of a longitude effect is in the frequency of magnetic bays. Dessler[16] notes that these are particularly frequent near Capetown, and that this may be due to the large electron-precipitation there.

One of the original reasons for this work was a hope of finding a longitudinal dependence of auroral protons, but as discussed in the last section this now seems most unlikely. This could have explained the unexpected rarity of auroral hydrogen emission seen from Invercargill, New Zealand (strong emission only once in the I.G.Y.-I.G.C. period (Gadsden, private communication); in contrast to observations from Yerkes Observatory[17] which is slightly further from the auroral zone).

Note—Professor J. A. van Allen (private communication) reports that outer-zone radiation at 1000 km altitude undergoes order of magnitude fluctuations from day to day, this being found from measurements with Explorer VII (1959ι). This suggests that the radiation regularly passes through the oxygen atmosphere within a day, rapidly enough to raise doubts about a longitude effect. But, if the electrons do traverse the oxygen atmosphere too rapidly, the lower intensity contours of the belt should not follow constant-field-strength contours; which, it seems, they are observed to follow.

Acknowledgments—This work is part of the research programme of the Dominion Physical Laboratory Auroral Station. The author is indebted to Mr. R. S. Unwin, Dr. M. Gadsden and Mr. W. H. Ward, in particular, for many helpful discussions.

REFERENCES

1. J. A. van Allen, *J. Geophys. Res.* **64**, 1683 (1959).
2. T. G. Northrop and E. Teller, *Phys. Rev.* **117**, 215 (1960).
3. R. L. Arnoldy, R. A. Hoffman and J. R. Winckler, *J. Geophys. Res.* **65**, 1361 (1960).
4. E. H. Vestine and W. L. Sibley *J. Geophys. Res.* **65**, 1967 (1960).
5. M. Nicolet *Ann. Géophys.* **15**, 1 (1959).
6. H. K. Kallmann, *J. Geophys. Res.* **64**, 615 (1959).
7. J. A. van Allen, C. E. McIlwain and G. H. Ludwig, *J. Geophys. Res.* **64**, 271 (1959).
8. H. Alfvén, *Cosmical Electrodynamics*, Chap. 2. Clarendon Press, Oxford (1950).
9. F. S. Johnson, *J. Geophys. Res.* **65**, 577 (1960).
10. E. C. Ray, *J. Geophys. Res.* **65**, 1125 (1960).
11. J. A. Welch, Jr. and W. A. Whittaker, *J. Geophys. Res.* **64**, 909 (1959).
12. E. Fermi, *Nuclear Physics*, Chicago University Press, Chicago (1950).
13. W. L. Fite, R. F. Stebbings, D. G. Hummer and R. T. Brackmann, *Phys. Rev.* **119**, 663 (1960).
14. A. J. Dessler and R. Karplus, *Phys. Rev. Lett.* **4**, 271 (1960).
15. P. Rothwell and C. E. McIlwain, *J. Geophys. Res.* **65**, 799 (1960).
16. A. J. Dessler, *J. Geophys. Res.* **64**, 714 (1959).
17. A. B. Meinel, *Proc. Nat. Acad. Sci., Wash.* **40**, 943 (1954).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 23–29. Printed in Great Britain.

A NEW INTERPRETATION OF THE STRUCTURE AND CO_2 CONTENT OF THE VENUS ATMOSPHERE

LEWIS D. KAPLAN*
Department of Meteorology, Massachusetts Institute of Technology, Cambridge, Mass.

(*Received* 2 *May* 1961)

Abstract—The observations of Venus in the near (Kuiper, Chamberlain and Kuiper) and far (Strong and Sinton) infrared are critically reviewed and interpreted in the light of recent laboratory determinations of the absorption of CO_2. These, together with the determination of the scale height at about 70 km above the cloud surface by observing the occultation of Regulus (de Vaucouleurs and Menzel), lead to a model of the Venus atmosphere and cloud layer temperatures that are self-consistent. It is concluded that a cloud-top temperature of 235°K, a cloud-top pressure of about 90 mB, and a CO_2 concentration of 15 per cent by volume best fit the observations. The amount of CO_2 above the cloud tops is then about 100 meters STP. The effective reflecting level for near infrared radiation is lower in the atmosphere, with about 400 meters STP of CO_2 above it. These lower clouds cannot be composed of water (or ice) particles.

1. INTRODUCTION

The nature of a planetary atmosphere is investigated by means of the radiation received from it. In the ultraviolet, visible and near-infrared portions of the spectra, most of the information comes from the reflected and scattered sunlight. If absorption bands occur, the presence of gaseous constituents can be determined, and careful photometry can sometimes put limits on their amounts. The limits can be narrowed greatly if an independent estimate is made of the pressure at some height above the reflecting surface, say by measurement of refraction of the light of the sun or, better, a star by the planetary atmosphere. Much information can often be obtained about the nature of the reflecting surface by visual observation and polarization measurements. In the far-infrared and microwave portions of the spectra, the information comes from the thermal emission of the planet itself and of its clouds and atmosphere. Careful interpretation of good measurements in absorption bands can, in principle, give information about the thermal structure of the planetary atmosphere and the temperature of the visible surface[1].

No attempt will be made here to give a complete review of even the best observations of Venus, and indeed many important clues will be disregarded because of the difficulty of their interpretation. The main purpose of this paper is to use a few key measurements to try to estimate the cloud-top temperature and put limits on the cloud-top pressure and the CO_2 concentration. For more completeness, a useful review of the Venus literature has been given by Shaw and Brobrovnikoff[2], and Mayer[3] has given a summary of the exciting recent microwave measurements.

The outstanding facts about Venus are that its surface is obscured by an almost uniform cloud or haze layer, of unknown thickness, and that it contains a large amount of CO_2. There is very little quantitative data on near-infrared CO_2-band photometry in the literature; and most of what little is usable of this is reported by Kuiper and his associates[4–7]. These few data, together with the emission measurements in the 8 to 13 micron region by Strong and Sinton[8, 9] and the recent interpretation by de Vaucouleurs and Menzel[10, 11] of the observations of the occultation of Regulus will be the basis of this study.

* Consultant to The RAND Corporation.

2. THEORY*

For the near-infrared bands of interest in this study, the spectrum is composed of lines the shape of whose absorption coefficients k_ν, to an accuracy measurable at interplanetary distances, are determined entirely by collisions. For a mixture of gases with fractions by volume x_i, the classical Lorentz formula for such lines is

$$k_\nu = \frac{S}{\pi} \frac{p\Sigma x_i \gamma_{is}}{(\nu - \nu_0)^2 + (p\Sigma x_i \gamma_{is})^2} \tag{1}$$

where ν_0 is the frequency of the line center, $S(T)$ the line intensity, and γ_{is} the line half-width for collisions with the gas i in the mixture at pressure $p=1$. A small dependence on temperature T of γ_{is} will be ignored.

The integrated absorption by such a line in a homogeneous path of reduced-length m is given by

$$\begin{aligned} \int_{line} A d\nu &= \int_{-\infty}^{+\infty} (1 - e^{-k_\nu m})\, d(\nu - \nu_0) \\ &= Sme^{-x} [I_0(x) + I_1(x)] \end{aligned} \tag{2}$$

where

$$x = \frac{Sm}{2\pi p \Sigma x_i \gamma_{is}} \tag{3}$$

and I_0 and I_1 are Bessel functions of imaginary arguments. This has the asymptotic form

$$\int_{line} A d\nu \longrightarrow 2\sqrt{Smp\Sigma x_i \gamma_{is}} \text{ as } x \longrightarrow \infty \tag{4}$$

For CO_2 lines detectable at interplanetary distances, equation (4) becomes an equality for p of the order of a tenth atmosphere or less. If, at the same time, the product Smp is small enough that the overlapping of adjacent lines is small, the absorption by an entire band or part of a band has a similar dependence on mp

$$\int_{band} A d\nu \doteq a\sqrt{mp}\,(1 + 0{\cdot}3\,x) \tag{5}$$

where x is now the CO_2 fraction by volume, and where a is a constant that can be determined by laboratory measurements. A CO_2–N_2 mixture is assumed in equation (5); Kostkowski[13] has found that γ_{is} is about 1·3 times as large for CO_2–CO_2 collisions as for CO_2–N_2 collisions.

The observed absorption of the CO_2 in Venus is small enough that equation (5) should be valid, within the accuracy of the measurements, providing the mean pressure of the "reflecting surface" is not much more than a fifth of an atmosphere.

For a path through a planetary atmosphere a mass-weighted mean $\bar{p}$ of the pressure should be taken. If the reasonable assumption is made that the CO_2 fraction by volume x is constant, this mean is one-half of the pressure of the reflecting surface p_{rs} (see discussion of Curtis Approximation in reference[12]). Thus, the absorption by a CO_2 band in the Venus atmosphere at an average angle θ with the vertical should be given by

$$\int_{band} A d\nu = a\sqrt{mp_{rs}}\,(1 + 0{\cdot}3\,x) \sec\theta \tag{6}$$

where m is the amount in the vertical. The factor of two due to the double pass through the atmosphere cancelled that in the pressure averaging.

3. INTERPRETATION OF ABSORPTION AND EMISSION MEASUREMENTS

Kuiper and Herzberg[1, 7] have attempted to estimate the amount of CO_2 in the Venus atmosphere by comparing the absorption with that obtained through long paths in the laboratory. Herzberg[7] reports that the strong bands near 1·6 μ are roughly matched by 88 m-atm of CO_2 in the laboratory at 1 atm pressure, but that more than 1400 m-atm of CO_2 are needed to match the weak bands near 1·05 μ. He attributes this discrepancy to re-emission at 1·6 μ, the different effect of scattering for strong and weak bands, and the effect of pressure-broadening. This writer is inclined to accept the last-mentioned effect as the major one, and believes that it is mostly due to the fact that the laboratory measurements were made at 1 atm, which is an order of magnitude higher than the mean pressures we will later derive for the atmosphere above the "reflecting level." The re-emission at 1·6 μ, in particular,

* See Reference[12] for a more complete discussion of some of the material in this section.

should be completely negligible compared with the reflected solar radiation.

The absorption by the 1·6 μ bands can be compared with the careful laboratory measurements by Howard, Burch, and Williams[14], whose data fit equation (5) almost exactly, with $a = 8{\cdot}4$ cm^{-1} if m is in m–atm, p in atm, and $\int Ad\nu$ in cm^{-1}.

For a path length of 88 m pure CO_2 and a pressure of 1 atm, the absorption should then be $8{\cdot}4\sqrt{88 \times 1{\cdot}3} = 90$ cm^{-1}. But the absorption obtained from the reproduced Venus spectra of Kuiper[6] is not less than 150 cm^{-1}, corresponding to a discrepancy of almost a factor of three in the product of the pressure and the path length.

The writer regards the measurements of Howard *et al.* as highly reliable, and is inclined to accept them as the basis of his computations. If the value of 8·4 is accepted for a and the average Venus air-mass sec θ is taken as 2, the pressure of an "effective reflecting surface" p_{re} can be computed from equation (6).

$$mp_{re} = \frac{(150)^2}{2\,(8{\cdot}4)^2\,(1 + 0{\cdot}3\,\alpha)} = \frac{159}{1 + 0{\cdot}3\,\alpha} \qquad (7)$$

If a value of 870 cm/sec^2 is accepted for the mean gravity, $m = 2{\cdot}61 \times 10^5\,\alpha p_{re}/M$ meters, where $M = 28{\cdot}02 + 15{\cdot}99\,\alpha$ is the molecular weight. From equation (7),

$$p_{re} = \frac{0{\cdot}0247\sqrt{M}}{\sqrt{\alpha + 0{\cdot}3\,\alpha^2}} \qquad (8)$$

Since the maximum value of α is 1, $p_{re} \geq 0{\cdot}14$ atmospheres.

The most recent opinion seems to be that the visible cloud top is the effective reflecting surface. Kuiper[15], for example, argues that the low emission temperature of about 235°K found by Strong and Sinton[8, 9] in the 8 to 13 μ region is due to strong molecular absorption and refers to layers well above the visible cloud, while the high rotational temperature of about 285°K found by Chamberlain and Kuiper[5] refers to a lower layer immediately above the cloud top. Since CO_2 is transparent between 8 and 9 μ, however, Kuiper is forced to postulate another strong molecular absorbent, such as carbon suboxide. The cloud-top temperature derived by Kuiper is 320°K, which implies a very high atmospheric opacity.

The opacity for this model of the Venus atmosphere can be estimated from equation (7), since the intensities of the 10 μ bands have been measured by Kostkowski[13]. At 285°K their intensity is less than that of the 1·6 μ bands as determined from the Howard, Burch, and Williams data[14]. Even considering that they are spread over a spectral interval one-third as wide as that of the 1·6 micron bands, the absorption from 9 to 11 μ should be under 60 per cent.

An emission corresponding to 235°K would therefore require unreasonably low upper air temperatures; and it must be concluded that the 235°K refers to the temperature at some level close to the top of the visible clouds. The author has recalculated the Venus emission in the 8 to 12 μ region, re-estimating the correction for absorption in the earth's atmosphere with the aid of lunar grating spectra kindly furnished him by Dr. Strong. The result is a uniform emission temperature of 225 ± 5°K throughout the entire range. As with the published results, the 10 μ bands show up only weakly. This is to be expected, as the bands are very temperature-dependent, decreasing in intensity by a factor of four for a temperature decrease from 285°K to 225°K. Even for an adiabatic lapse rate, the pressure of the cloud top would be not more than a half of that for which the above absorption estimates for the band at 1·6 μ were made. The author estimates that the *maximum* absorption in the 9·4 μ band by the layer above the clouds is less than 30 per cent.

Since the emissivity of the clouds in the 8 to 10 μ region can be expected to be somewhat less than unity, 235°K seems to be a reasonable value for the cloud-top temperature.

The difference between the rotational and emission temperatures can be attributed to a high opacity of the cloud particles around 10 μ and a high transparency in the near-infrared. It is interesting, in this connection, that comparison of the radiation received in the window just beyond the 1·6 μ bands with that at shorter wavelengths indicates that the spectral shape of the scattered light in the absence of absorp-

tion is the same as that of the incident sunlight[6]. This is what would be expected from a thick cloud consisting of large transparent particles.

It should be noted that the foregoing discussion does not necessarily invalidate Kuiper's argument that the clouds cannot be composed of water droplets. On the contrary, the long path through a saturated water layer should produce detectable water bands in the visible and near-infrared.

The low cloud-top temperature implies an albedo of the order of 80 per cent which means that the ground surface of Venus must receive much less solar radiation than the earth does. In order to provide a large enough greenhouse effect to maintain the high surface temperatures determined by the radio measurements, the atmospheric transparency beyond $4\,\mu$ would have to be very small. On the other hand, the cloud particles would have to be transparent enough to sunlight to allow a substantial portion of the remaining solar radiation to reach the ground. Also the clouds could not be more than a few kilometers thick.

Sagan[16] has shown that the CO_2 absorption is insufficient to account for the required atmospheric opacity, and finds that sufficient water vapor to account for the clouds as ice-crystals will serve the purpose. Unpublished measurements by Strong suggest that this is more water-vapor than Venus has; but Sagan's hypothesis is certainly attractive enough to warrant further investigation.

Part of the radiation received from Venus at small wavelengths would have made a double traversal through the entire atmosphere.

If Chamberlain and Kuiper's value of 285° is taken for the temperature at a level having half the pressure of the effective reflecting surface, and an adiabatic lapse-rate is assumed between it and the cloud top at 235°K, the pressure p_c at the cloud top will be about one-fourth of p_{rs}. From equation (8), therefore,

$$p_c \doteq 0{\cdot}033 \sqrt{\frac{1+0{\cdot}57\,\alpha}{\alpha+0{\cdot}3\,\alpha^2}}\ \text{atm.} \qquad (9)$$

This is presented by the dashed curve in Fig. 1.

If the microwave measurements (giving about 585°K) are accepted as representing thermal radiation from the surface, the effective reflecting surface would have to be a second cloud level below that of the visible cloud tops. These lower clouds cannot be composed of water particles.

4. INTERPRETATION OF REFRACTION MEASUREMENTS

From measurements of the intensity and apparent position of Regulus during its occultation by Venus on 7 July 1959, Menzel and de Vaucouleurs[11] were able to estimate the scale height $H = RT/Mg$, its derivative $\delta H/\delta Z$ and the pressure at the altitude $Z_0 = 70 \pm 8$ km above the top of the visible cloud layer that corresponds to half-intensity. The results reported were $H(Z_0) = 6{\cdot}8 \pm 0{\cdot}2$ km:

$$\delta \ln H/\delta Z = 0{\cdot}010 \pm 0{\cdot}002\ \text{km}^{-1};$$

and $p(Z_0) = 0{\cdot}0026 \pm 0{\cdot}00013$ mB.

The method used to obtain these results is described by de Vaucouleurs[10], and involves an assumption of an isothermal atmosphere of constant composition above the level at which H and p are determined. It should not be too difficult to determine the effect on the results of an increase in temperature and decrease of molecular weight with height.

The composition chosen by Menzel and de Vaucouleurs was 90 per cent CO_2, 9 per cent N_2 and 1 per cent A. The molecular weight M is then 42·5. For our mixture of α CO_2 and $(1-\alpha)$ N_2

$$M = 28{\cdot}02 + 15{\cdot}99\,\alpha \qquad (10)$$

and

$$T(Z_0) = MgH/R = 195{\cdot}8 + 111{\cdot}7\,\alpha\ ^\circ\text{K} \qquad (11)$$

M and $T(Z_0)$ are given in Table 1 as functions of α, assuming $g = 860$ cm sec^{-2}. A significant portion of $\delta H/\delta Z$ at Z_0 should be due to a decrease of M with height due to photochemical decomposition of CO_2. Also, for a given value of α for the lower atmosphere, $T(Z_0)$ should be somewhat lower than the value given at that α in Table 1.

With Z_0 taken as 70 km, the pressure p_c at the visible cloud top can be determined from the hydrostatic equation as a function of the mean temperature and the CO_2 concentration.

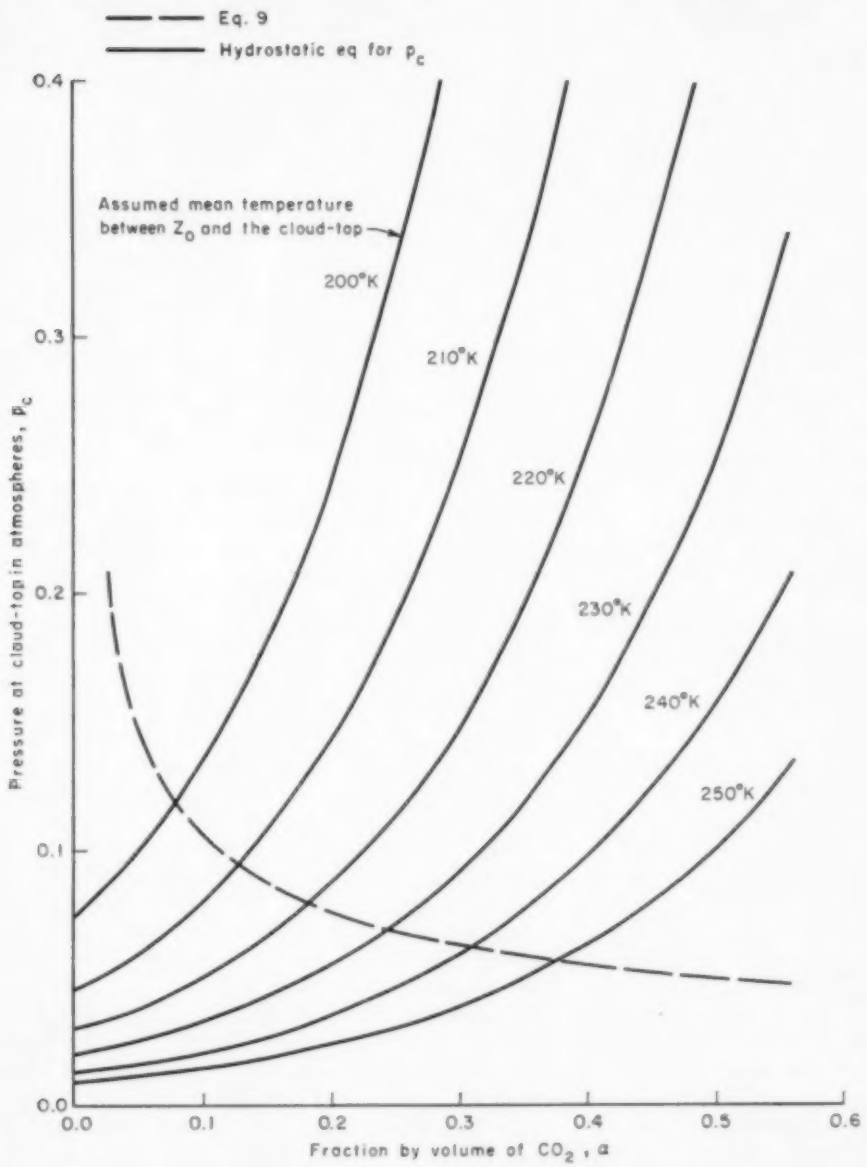


Fig. 1. The pressure at the Venus visible cloud-top, p_c, as a function of the fraction of $C0_2$ by volume, α. The dashed line (from equation (9)) is based on an interpretation of infrared observations; the solid lines are based on an interpretation of the measurements of the occultation of Regulus for various values of the mean temperature between the visible cloud-top and the level at which the occultation occurred.

Table 1. Molecular weight and temperature at half-intensity height as functions of CO_2 concentration

α	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0
M	29·62	31·22	32·82	34·42	36·02	37·61	39·21	40·81	42·41	44·01
$T(Z_0)$°K	207	218	229	241	252	263	274	285	296	308

This was done for several values of mean temperature, and the results are given by the solid curves in Fig. 1. The acceleration of gravity was taken as $g=870$ cm sec^{-2}.

If the mean temperature can be estimated, the CO_2 concentration and the pressure of the visible cloud top are given by the point of intersection with the dashed curve. If it is assumed that the temperature distribution above the cloud top is not far from that which would obtain at radiative equilibrium, the temperature must decrease with height to reach a minimum somewhat below 190°K. The temperature at its maximum due to ozone will depend on the O_2 concentration, but will almost certainly be considerably lower than that observed in the earth's atmosphere. It follows, therefore, from Table 1 and Fig. 1 that the CO_2 concentration must be considerably less than 50 per cent. A reasonable guess at the mean temperature seems to be about 210 to 220°K, which yields a concentration $\alpha=0{\cdot}15$ and a visible cloud-top pressure $p_c=90$ mB.

These quantities are probably not reliable to better than 50 per cent. The major source of error seems to be in the estimate of the distance Z_0 between the half-intensity height and the height of the visible cloud-top. According to a private communication from Dr. de Vaucouleurs, the major part of the 8 km uncertainty in Z_0 is in the determination of the position of the visible cloud-top. It would seem that much of this uncertainty can be removed by careful measurement of existing plates.

If Herzberg's numbers are accepted rather than Howard's, or if the temperature gradient below the upper cloud level is less than adiabatic, both the CO_2 concentration and the pressure of the visible cloud-top have to be reduced.

5. CONCLUDING REMARKS

The principal results obtained in this paper are a cloud-top temperature of about 235°K, a cloud-top pressure of about 90 mB, and a CO_2 concentration of about 15 per cent by volume. The corresponding CO_2 amounts are about 100 meters STP above the cloud top and 400 meters STP above the "effective reflecting level." This is the first time, to the author's knowledge, that a self-consistent estimate has been made of the cloud-top pressure and the CO_2 concentration separately.

If a temperature of 580°K is accepted for the Venus surface and a lapse-rate close to adiabatic is assumed, the surface pressure is of the order of two atmospheres and the total CO_2 of the order of 2 km STP.

The above results obviously cannot be taken as definitive. They are based on many assumptions, which, though they seem reasonable to the author are certainly not obvious. Moreover, barely enough relationships were used to get the results obtained.

A further check on the consistency of the results could come from a theoretical investigation of the photochemistry of an atmosphere having the derived composition. It would also be a great help to have more of the Venus spectral measurements and visual observations published in a quantitative form.

The author hopes to have at least provided a model of the Venus atmosphere that is self-consistent and unique within the framework of the assumptions. He feels that more information is available in the observations than has yet been extracted from them, but that better, carefully planned observations are necessary to obtain a convincing picture of the Venus atmosphere and a satisfactory explanation of its thermal structure. The energy limitation of earth-bound spectroscopic measurements probably requires that a major part of the future measurements be made from close-approach vehicles.

Acknowledgements—The number of people with whom the author has had helpful discussions concerning this paper is too large to allow mentioning all their names here. He is, however, particularly indebted to Professors de Vaucouleurs and Öpik.

REFERENCES

1. L. D. KAPLAN, *J. Opt. Soc. Amer.* **49,** 1004 (1959).
2. J. H. SHAW and N. T. BOBROVNIKOFF, *Natural Environment of the Planet Venus*, WADC Phase Technical Note 847-2, February 1959.
3. C. H. MAYER, *Astr. J.* **64,** 43 (1959).
4. G. P. KUIPER (ed.), *The Atmospheres of the Earth and Planets*, pp. 306, 406 (2nd Ed.). University of Chicago Press, Chicago (1952).
5. J. W. CHAMBERLAIN and G. P. KUIPER, *Astrophys. J.* **125,** 399 (1956).
6. G. P. KUIPER, *Astrophys. J.* **106,** pp. 248, 252 (1947).
7. G. HERZBERG, *J. R. Astr. Soc. Can.* **45,** 100 (1951).
8. W. M. SINTON, Doctoral Dissertation, Johns Hopkins University, 1956.
9. J. STRONG and W. M. SINTON, *Astrophys. J.* **131,** 460 (1960).
10. G. DE VAUCOULEURS, Air Force Missile Development Center Technical Note (AFMDC-TN-59-37), Holloman A. F. Base, New Mexico, December 1959.
11. G. DE VAUCOULEURS and D. H. MENZEL, *Nature, Lond.* **188,** 28 (1960).
12. L. D. KAPLAN, *The Atmosphere and the Sea in Motion*, B. Bolin (ed.), p. 170. Rockefeller Institute Press and Oxford University Press, New York (1959).
13. H. J. KOSTKOWSKI, Doctoral Dissertation, Johns Hopkins Univiersity, 1954.
14. J. N. HOWARD, D. L. BURCH and D. WILLIAMS, *G.R.D. Paper No.* 40, Air Force Cambridge Research Center, November 1955.
15. G. P. KUIPER, *The Threshold of Space*, M. Zelikoff (ed.), p. 78. Pergamon Press, London (1957).
16. C. SAGAN, Tech. Report No. 32-34, Jet Propulsion Laboratory.

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 30-34. Printed in Great Britain.

THE VIABILITY OF MICRO-ORGANISMS IN ULTRA-HIGH VACUUM

E. E. BRUESCHKE, R. H. SUESS and M. WILLARD
Hughes Aircraft Company, Aerospace Group, Culver City, California

(*Received* 4 *April* 1961; *revised manuscript received* 28 *May* 1961)

Abstract—Studies have been conducted on the resistance of common micro-organisms to vacuum from 1×10^{-6} to 6×10^{-9} mm Hg. Some of the organisms remained viable after exposure to 8×10^{-9} mm Hg. All were killed by prolonged exposure to pressures of $1{\cdot}2 \times 10^{-9}$ mm Hg and lower.

1. INTRODUCTION

There has been considerable speculation about whether or not spores and similar life forms can survive the rigors of travel through interplanetary space. These spores may be pushed through space by light pressure[1] or may be present on space vehicle surfaces due to contamination subsequent to sterilization of such vehicles. As part of studies on surface sterilization problems for such projects as the lunar Surveyor spacecraft, the survival of spores at extremely low pressures has been studied.

Previous investigations demonstrate that micro-organisms can survive moderately reduced pressures[2]. Proctor and Parker[3] have collected viable micro-organisms at an altitude of 36,000 feet using the stratosphere balloon Explorer II. In addition, Flosdorf and Mudd[4] have used vacuum extensively as part of a micro-organism preservation technique. Their method is to evaporate a broth of micro-organisms in a desiccator containing phosphorus pentoxide. The tubes are then sealed off in a vacuum of approximately 10^{-2} mm Hg and can be stored for years in this condition. One investigator, Bakanauskas[5], exposed spores of *Bacillus globigii* (*subtilis*), *Bacillus mycoides, Bacillus cereus, Aspergillus flavus, and Aspergillus niger* to pressures as low as 5×10^{-7} mm Hg for extended periods. All of the micro-organisms survived except *Bacillus cereus.*

The pressure at approximately 325 miles altitude is 10^{-6} mm Hg and the pressure in outer space is less than 10^{-12} mm Hg. These reduced pressures present beyond the earth's atmosphere suggested a need for the study of the viability of spores subjected to ultra-high vacuum. Thus, the resistance of four common micro-organisms to pressures of 1×10^{-6} to 6×10^{-9} mm Hg was investigated in this laboratory. The following organisms were selected for evaluation:

1. *Bacillus subtilis*
2. *Aspergillus niger*
3. *Aspergillus terreus*
4. *Penicillium citrinum*

The first is a bacterium which exhibits a close physiological relationship to pathogens. The remainder are fungi which commonly attack materials.

2. EXPERIMENTAL TECHNIQUES AND TEST PROCEDURES

The experiment was carried out in the apparatus shown in Figs. 1 and 2. The specimen tube assembly (Fig. 1) consisted of four specimen tubes joined to a central manifold. Each tube had a side arm for admitting sterile air at the conclusion of the test. The specimen tube assembly was sealed to the high vacuum system described below.

Description and operation of high vacuum system

The ultra-high vacuum system shown in Fig. 2 consisted of a 75 l/s mercury diffusion pump and an 8 l/s ion pump. A Dewar flask containing a dry ice-acetone mixture was used at cold trap No. 2 to condense water and other volatiles evaporated from the specimen manifold. Initial pumping was accomplished with

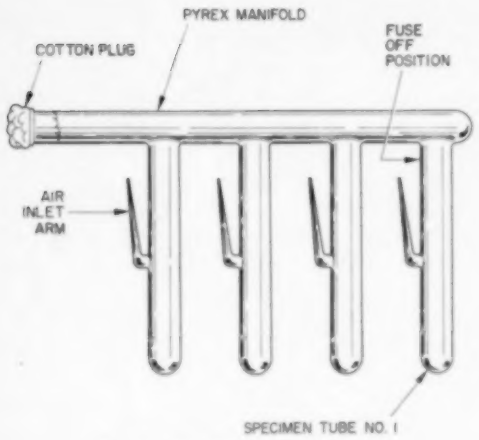
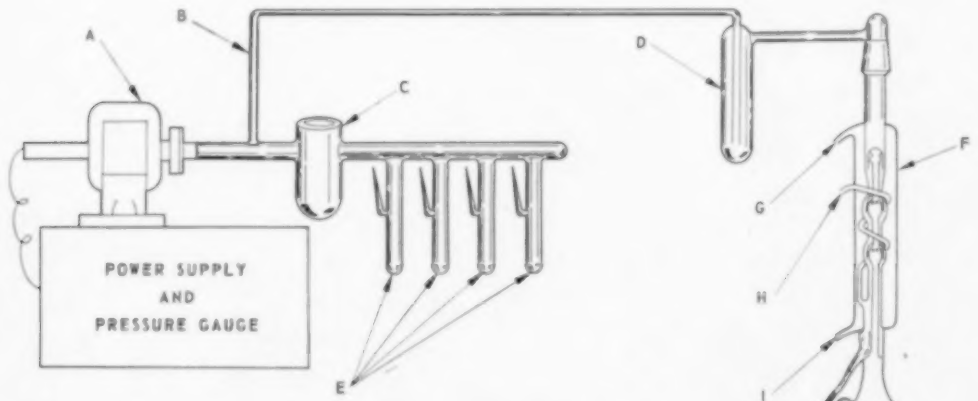


Fig. 1. Specimen tube assembly.

the mechanical pump alone. The mercury diffusion pump was started when the system pressure was about 10 microns. After 10^{-4} mm Hg was reached the ion pump was started. This procedure is necessary as ion pumps cannot operate at pressures greater than 10^{-2} mm Hg and should not be started until 10^{-4} mm Hg is reached. Following the initiation of pumping by the ion pump, the mercury diffusion system was removed by fusing off the Pyrex tube at point B. During fuse-off of the specimen tubes, the pressure in the system will increase due to gases being liberated from the inner surface of the glass. In these instances, liquid nitrogen is placed in cold trap No. 1 to insure that the pressure in the system does not rise high enough to cause the ion pump to cease pumping.

Chamber pressure was determined by measuring the ion pump current. The ion current versus pressure relationship for the pump used was provided by the manufacturer and calibrated in the laboratory using a precision McLeod gauge.

Exposure of micro-organisms to high vacuum. A paper strip containing spores of *B. subtilis* was aseptically placed in a sterile tube containing 3 ml of distilled water. The tube was shaken to disperse the spores, and the strip was removed with sterile tweezers. Three ml sterile water was added to the culture tube of each fungus and the tubes were shaken to disperse the spores. All the spore suspensions were then aseptically combined and transferred to the sterile specimen tube assembly. The



A ION PUMP
B SEAL OFF POINT
C COLD TRAP #1
D COLD TRAP #2
E SPECIMEN TUBES
F 3 STAGE MERCURY DIFFUSION PUMP
G COOLING WATER OUTLET
H TO ROUGHING PUMP
I COOLING WATER INLET
J ELECTRICAL HEATER

Fig. 2. Ultra-high vacuum system.

assembly was sealed to the high vacuum system and the water was evaporated at room temperature. The pressure was then reduced to 10^{-4} mm Hg using the mercury diffusion pump and to lower pressures using the ion pump. A pressure of 8×10^{-8} mm Hg was reached after ten days of pumping. The first specimen tube was then fused off from the apparatus and checked for spore viability. The schedule for removal of the specimen tubes is given below and shown in Fig. 3.

Culture tube	*Days exposure*	*Pressure at end of exposure period, mm Hg*
1	10	8×10^{-5}
2	30	$1{\cdot}2 \times 10^{-8}$
3	40	6×10^{-9}
4	45	6×10^{-9}

Each sample tube was sealed off from the manifold in such a manner as to maintain the reduced pressure in both the sample tube and the main system.

After removal the side arm of the sample tube was scored with a file, flame sterilized, and attached to a sterilized air filter (Fig. 4) consisting of a 0·5 μ millipore filter, absorbent cotton and activated charcoal. The tip was then broken within the rubber connector and filtered air allowed into the sample tube.

Test for viability. Viability of the micro-organisms after exposure was determined by the streak plate method. Sterile distilled water was introduced into the specimen tube and the tube was carefully shaken to disperse the spores. 0·1 ml portions of the suspension were withdrawn with a sterile pipet, planted on nutrient agar and wort agar media, and streaked. The two types of media were used to insure optimum culture conditions for every micro-organism. The inoculated media were incubated at 33°C for ten days, at the end of which time they were examined for growth. Colonies of fungi were examined on the agar medium at 430 magnification. A stained slide

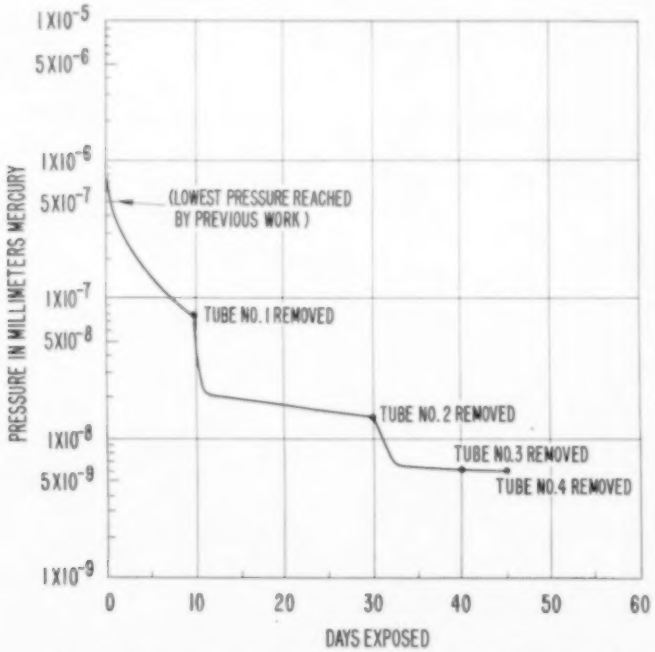


Fig. 3. Exposure schedule of micro-organisms subjected to ultra-high vacuum. Previous work cited is that of Bakanauskas[5].

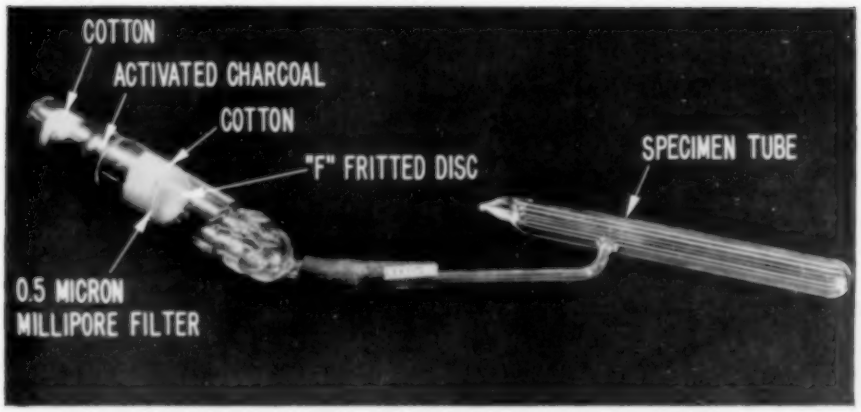


Fig. 4. Apparatus for admitting sterile air into specimen tube.

(*To face page* 32)

of any bacteria colonies growing was prepared and examined at 970 magnification.

Viability control. Suspensions of the unexposed spores of each micro-organism tested were planted on nutrient agar and wort agar media at the start and at the end of the test period. In addition, a viability control test was conducted in which a suspension of spores, prepared as in the test samples, was dried and rapidly brought to a pressure of 10^{-6} mm Hg. The spores were then resuspended in distilled water and 0·1 ml portions planted on nutrient and wort agar. The inoculated media were incubated at 33°C for ten days, at the end of which time they were examined for growth.

3. RESULTS

After two days incubation colonies of *B. subtilus*, *A. niger*, and *A. terreus* were observed on agar inoculated with spore suspension from specimen tube No. 1. The colonies appeared normal although fewer in number than on the control agar. Colonies of *P. citrinum* were not found.

No viability of any of the micro-organisms was observed after exposure to pressures of $1{\cdot}2 \times 10^{-8}$ to 6×10^{-9} mm Hg (tubes 2, 3, and 4).

To determine whether the reduction in pressure to $1{\cdot}2 \times 10^{-8}$ mm Hg and lower, or the longer exposure of the tubes subjected to these pressures had been responsible for the destruction of the micro-organisms, another exposure test was conducted. The pressure of a specimen tube of the micro-organisms was reduced to 8×10^{-9} over a period of only ten days. The viability test after this exposure showed that the three micro-organisms, *B. subtilis*, *A. niger*, and *A. terreus*, survived.

4. CONCLUSIONS

Exposure of the micro-organism spores tested to pressures of $1{\cdot}2 \times 10^{-8}$ mm Hg and lower for a sufficient period of time will cause their destruction. This time for all the micro-organisms except *P. citrinum* is greater than ten days but may be less than thirty days. Further effort must be expended to determine the minimum exposure time for lethal action of high vacuum on various families of micro-organisms.

Chemical analysis indicates no significant difference in total water content of spores and vegetative cells[6]. It has also been shown that spores possess an active metabolism and that they contain a full complement of the same kinds of enzymes and coenzymes found in vegetative cells but in much lower quantities. Thus, the mechanism of destruction of spores by ultra-high vacuum may be the evaporation of cellular water, resulting in crenation and fracture of the cell wall. The time dependency of destruction by ultra-high vacuum suggests other mechanisms of annihilation whereby oxygen is depleted and intermediate products are accumulated. These factors can be studied by determining wet weight and dry weight differences on the spores after vacuum exposure and by conducting experiments on anaerobic micro-organisms. It is possible that the vapor pressure of cellular constituents themselves may have been reached, resulting in their gradual sublimation in ultra-high vacuum.

The results of this experiment imply that inadvertent surface contamination of space vehicles during pre-launch periods may not be a serious problem with space flights of long duration since the surface may be vacuum sterilized during flight. Even on a short flight of 63 hours to the moon, micro-organisms may be destroyed by the combination of ultraviolet radiation and high vacuum. If not, surviving organisms arriving at the moon may still be killed, in time, as its surface pressure is approximately 10^{-10} mm Hg. Conversely, organisms buried beneath the moon's surface on impact of the space vehicle may survive.

It appears that unprotected spores borne through interplanetary space by such methods as light pressure will be killed by the ultra-high vacuum of greater than 10^{-12} mm Hg present in space. If this is true, then this postulated mechanism by which life is propagated in the universe may be invalid.

REFERENCES

1. M. H. Briggs. *Spaceflight* 2, No. 2, April, 39 (1959).

2. W. W. C. TOPLEY and G. S. WILSON. *Principles of Bacteriology and Immunity* (3 ed.), Vol. I, Williams and Wilkins, Baltimore (1946).

3. B. E. PROCTOR and B. W. PARKER. *Aerobiology* (*Ed.* F. R. Moutton), Washington, D.C. (1942).

4. E. E. FLOSDORF and S. MUDD. *Journal of Immunol.*, **29**, 389 (1935).

5. S. BAKANAUSKAS. *Resistance of Microorganisms to High Vacuum, WADC TN* 59-142, September 1959.

6. C. LAMANNA. *Bact. Rev.*, **16**, 89 (1952).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 35–42. Printed in Great Britain.

DIFFUSION COEFFICIENTS DETERMINED FROM SODIUM VAPOUR TRAILS

J. A. REES*
Department of Applied Mathematics, Queen's University, Belfast

(*Received* 8 *May* 1961)

Abstract—The examination of photographs of a trail of sodium vapour laid by a Skylark rocket at latitude 31°S has led to the determination of the molecular diffusion coefficient for sodium at altitudes between about 110 and 140 km. The diffusion coefficients obtained are in general agreement with those found by Blamont and by Manring *et al.* by similar methods, but higher than those calculated from the 1959 A.R.D.C. Standard Atmosphere. Values of the atmospheric scale height and mass density for altitudes between 110 and 140 km are deduced.

1. INTRODUCTION

It was suggested by Bates[1] that useful information concerning upper atmospheric conditions might be obtained by increasing locally the amount of free sodium in the upper atmosphere. Following this suggestion four Skylark rockets carrying sodium burners designed by Armstrong have to date been launched from Woomera (lat. 31°S) for the group at Queen's University, Belfast. The rockets were launched during evening twilight and all successfully laid trails of sodium vapour between about 70 km and the tops of their trajectories (which varied between 130 and 220 km). The firings were primarily carried out to examine upper atmosphere winds and temperatures. The data on winds obtained from Skylark 11 (fired 3 December 1958) have been described by Groves[2], and the data obtained from Skylark 16 (fired 30 November 1959) by the two methods outlined by Rees[3] are being prepared for publication[4].

The photographs (taken from two or more camera sites) of the sodium trails used in the wind studies may also be used for a determination of the rate of diffusion of the sodium atoms. Fig. 1 consists of two photographs of the trail of Skylark 16, taken with the same camera, at an interval of six minutes. The photographs show clearly the growth of the diameter of the trail and also that the growth was more marked at the higher altitudes. Measurements of the rate of change of the diameter of the trail at various known altitudes have led to the determination of the diffusion coefficient, D, for sodium at these altitudes.

2. THEORETICAL ANALYSIS

A detailed analysis of the growth of sodium clouds released from rockets, either as discrete bursts or as continuous trails, has been given by Dalgarno[5]. The present analysis is a generalization of Dalgarno's treatment. It is supposed that the trail forms an infinitely long cylinder of circular cross-section.

After a short ($\leqslant 1$ sec) interval in which the sodium atoms reach thermal equilibrium with the atmospheric molecules, the most probable distribution of the number density, $n(r, t)$, of the sodium atoms at a perpendicular distance r from the axis of the cylinder and at time t is the Gaussian distribution

$$n(r, t) = (N/\pi R^2) \exp(-r^2/R^2) \qquad (1)$$

where

$$R^2 = 4Dt \qquad (2)$$

and N is the initial line density of sodium atoms along the trail axis. Equation (1) ignores the effect on $n(r, t)$ of the prevalent winds.

It follows that the number of sodium atoms in a unit column along a line of sight through the trail at a distance s from the trail axis is given by

$$S = (N/\pi^{\frac{1}{2}} R \sin\theta) \cdot \exp(-s^2/R^2) \qquad (3)$$

* Now at Institute of Advanced Studies, Australian National University, Canberra, Australia.

where θ is the angle the line of sight makes with the trail axis. If each atom radiates λ 5893 Å at a rate of p photons per sec, then neglecting self-absorption and multiple-scattering the photon intensity of the 5893 Å radiation arriving at a camera is given by

$$I_s = Sp = (Np/\pi^{\frac{1}{2}}R \sin\theta). \exp(-s^2/R^2) \text{ ph/cm}^2 \text{ col.} \quad (4)$$

so that

$$\exp(-s^2/R^2) = I_s/I_0 = 1/X \text{ (say)}$$

which using equation (2) becomes

$$s^2 = 4Dt \log_e(X). \quad (5)$$

Thus, if the intensity distribution across a diameter of the trail is examined, the half width s at an intensity

$$I_s = I_0/X$$

is given by equation (5); and if s is measured for a particular altitude and a given value of X, for various times t, the graph of s^2 against t is a straight line of slope $4D \log_e X$ from which D is obtained.

In the special case where

$$X = e (\equiv 2{\cdot}718\ldots)$$

equation (5) reduces to

$$s^2 = 4Dt. \quad (6)$$

The slope of the line obtained by plotting s^2 against t is then $4D$. This is the basis of the method used by Blamont[6] and by Manring *et al.*[7].

An alternative method of determining D, which to some extent reduces the errors introduced by neglecting self-absorption and multiple scattering, is as follows. From equation (5)

$$\log_{10} I_s = \{9{\cdot}212\, Dt \log_{10} I_0 - s^2\}/9{\cdot}212\, Dt$$

which becomes in terms of the density, B, of the photographic film,

$$B_s = \gamma\{9{\cdot}212\, Dt \log_{10} I_0 - s^2\}/9{\cdot}212\, Dt \quad (7)$$

where γ is a constant for the particular film. A graph of B_s against s^2 should, therefore, be the straight line PQ in Fig. 2. However, if self-absorption is important, its influence will

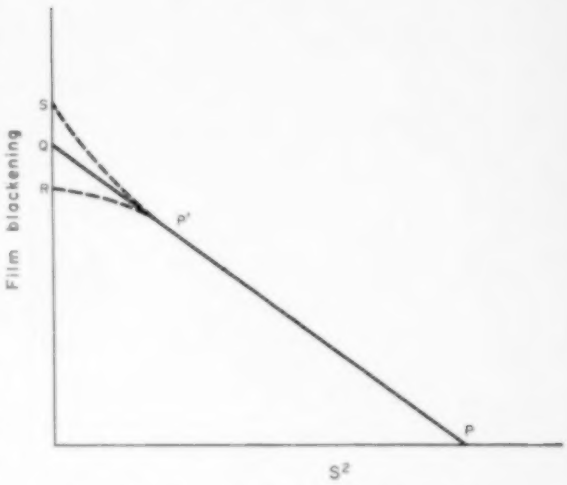


Fig. 2. Variation of film blackening with trail half-width.

be most marked at small values of s and the graph of B_s against s^2 will be PR and not PQ. If, on the other hand, an appreciable quantity of solid matter is ejected from the sodium burners, the graph may be of the form PS. The slope, G, of the straight line portion, PP', is given by

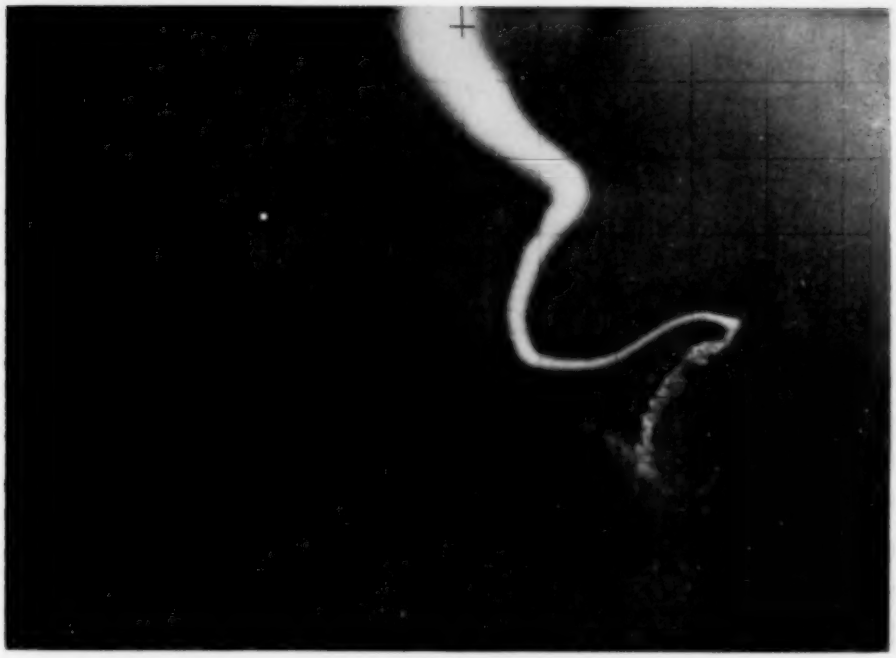
$$G = \gamma/9{\cdot}212\, Dt \quad (8)$$

from which Dt may be determined if γ is known for the film.

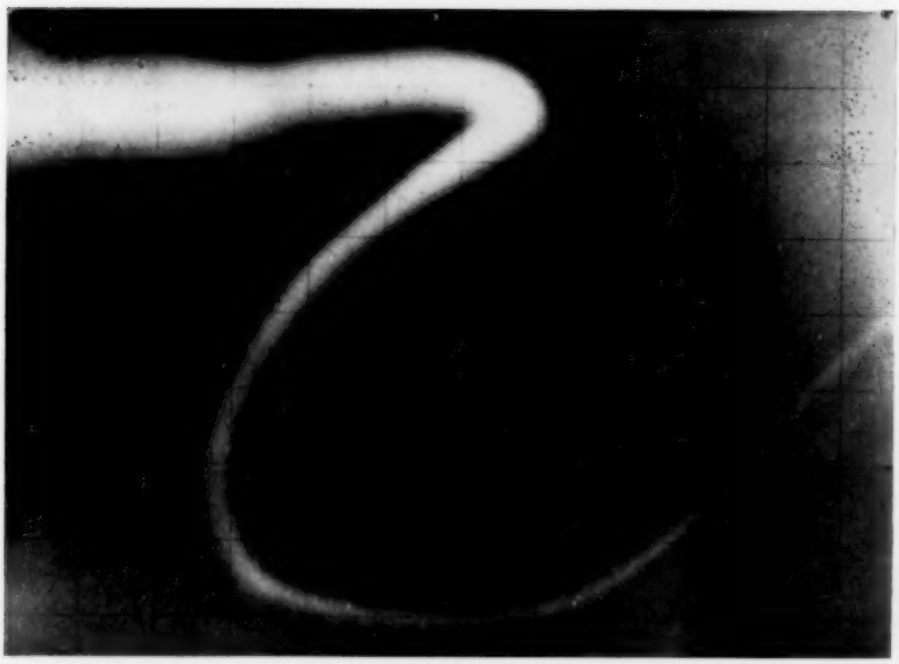
By plotting B_s against s^2 for a particular altitude and a range of times, t, a series of values of Dt may be obtained. Since the origin of the time scale is unknown, these values of Dt must be plotted against t, measured from an arbitrary zero, and the corresponding value of D found.

3. EXPERIMENTAL METHOD

The sodium from Skylark 16 was ejected in a continuous stream between about 75 km and 150 km during evening twilight. The trail was photographed from three camera positions distributed as shown in Fig. 3. The most suitable photographs for diffusion measurements were those taken from position C18 with a F24 camera (focal length 20 cm) using HP3 film and exposure times of up to 15 sec. Four photo-



(a)



(b)

Fig. 1. Photographs of sodium trail. (a) Approx. 3½ mins after formation; (b) Approx. 9½ mins after formation.

(*To face page* 36)

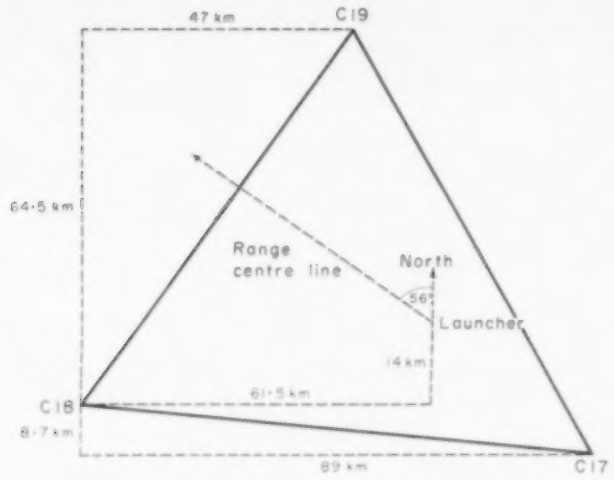


Fig. 3. Distribution of camera sites and "Skylark" launcher. C17, C18 and C19 are camera sites.

graphs taken over an interval of 8 minutes were selected, two of which are shown in Fig. 1. The photographs show the diameter of the trail to increase with time, the growth being most marked in the higher portions of the trail. It is seen from Fig. 1a that turbulence was very marked in the lower parts of the trail but stopped abruptly at an altitude of $103 \pm \frac{1}{2}$ km. Similar observations have been reported by other investigators: the value of 103 km found for the upper limit in the present work is in good agreement with the values of 102 km reported by Groves[2], 100 ± 1 km found by Blamont[6, 8], and 97 km found by Manring[9]. The trail below 103 km faded rapidly and disappeared within about ten minutes of its formation.

On the selected photographs, corresponding parts of the trail at the different times were first chosen and their positions in space, with respect to the camera at C18, determined by the Projector Method[3]. Ideally, if point P_1 on the photograph taken at time t_1 corresponds to point P_2 on the photograph taken at time t_2 and so on, the points $P_1, P_2, \ldots P_n$ should be at the same altitude with other series of corresponding points at other altitudes. However, if the prevalent winds at these altitudes have significant vertical components, $P_1, P_2, \ldots P_n$ will be at different altitudes and the diffusion coefficient found by comparing the half-widths of the trail at the points $P_1, P_2, \ldots P_n$ will be some average value. In the case of Skylark 16, the vertical component of the wind at altitudes between 110 and 140 km was such that the error in the altitude attributed to a particular diffusion coefficient was generally less than ± 1 km.

The intensity profiles of the sodium trail at a selected altitude (between 110 and 140 km) were obtained by using a Joyce-Loebl Microdensitometer to measure the variation of the density of the images of the trail in a direction at right angles to the trail axis. From these profiles the half widths, s, of the trail were determined at several values of X between 1·1 and 1·75 for a series of times, t. The values of s were of course suitably corrected to allow for the various distances of the portions of the trail from the camera. The values of s^2 so obtained were plotted against t and from these graphs the diffusion coefficient D was determined using equation (5). By taking D as the mean of the values found for several values of X, any errors due to irregularities in the profiles of the trail were reduced.

The diffusion coefficients for the various altitudes were also found by the alternative

method outlined above. The values of D found for a particular altitude by the two methods agreed to within 10 per cent and are shown in Fig. 4.

4. DISCUSSION OF RESULTS

As well as the diffusion coefficients found for Skylark 16, Fig. 4 shows the values determined by Blamont[10], and by Manring *et al.*[7,11] of G.C.A. from sodium trail experiments. The data refer to measurements made at very nearly the same latitude of 30°, but in different hemispheres. There is some measure of agreement between the various measurements, although at altitudes near 120 and 150 km there is a spread of a factor of three.

Also included in Fig. 4 are values of D calculated by the author from the 1959 A.R.D.C. Standard Atmosphere[12] values of atmospheric number density N, mean molecular mass M and temperature T using the formula

$$D = \{(3\pi^{\frac{1}{2}}/16N) \cdot (2kT/\mu)^{\frac{1}{2}}\}/Q_D \qquad (9)$$

with

$$\mu = mM/(m+M) \qquad (10)$$

k being Boltzman's constant, m the mass of a sodium atom and Q_D the cross-section for diffusion (Chapman and Cowling[13]). Unfortunately the appropriate value of Q_D is not accurately known. Following Shklovskii and Kurt[14], Q_D has been taken as $3{\cdot}85 \times 10^{-15}$ cm^2, but it should be remembered in comparing values of D calculated from equation (9) with those found by the rocket experiment that Q_D is somewhat uncertain, an error of 30 per cent being not impossible. It is seen from Fig. 4 that even allowing for such an error in Q_D, the diffusion coefficients calculated using equation (9) are different from those deduced from the observations.

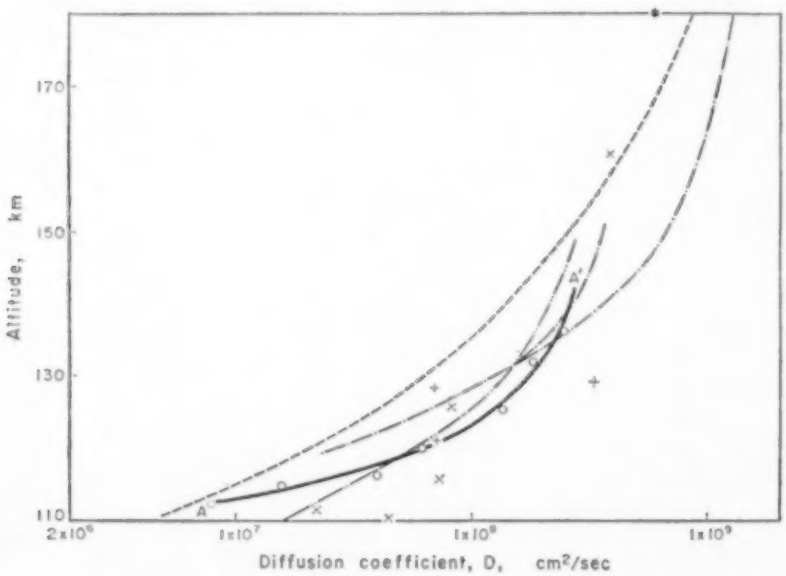


Fig. 4. Variation with altitude of diffusion co- efficients for sodium.
○"Skylark 16" data
— — — —ARDC 1959 Standard Atmosphere[12]
—·—·—·—Blamont[10]
—··—··—G.C.A. [11]
× Zimmerman and Champion[17]
+ Marmo et al. [15,16]
*Manring et al.[7]

In an analysis of their artificial electron-cloud experiments, Marmo *et al.*[15] found the diffusion coefficient for potassium at 121 km to be $6{\cdot}2\times10^{7}$ cm^2/sec and that for caesium at 128 km to be $2{\cdot}5\times10^{8}$ cm^2/sec. The corresponding coefficients for sodium have been deduced by allowing for the differences in the masses of the sodium, potassium and caesium atoms. The values found in this way are $7{\cdot}1\times10^{7}$ cm^2/sec at 121 km and $3{\cdot}4\times10^{8}$ cm^2/sec at 128 km. No allowance has been made for differences in the diffusion cross-section, Q_D for sodium, potassium and caesium. The diffusion coefficients deduced for sodium may, therefore, be in error by ~20 per cent. The value of $7{\cdot}1\times10^{7}$ cm^2/sec at 121 km is in excellent accord with the present measurements. The agreement at 128 km is poorer—but still better than that with the value calculated from the 1959 A.R.D.C. Standard Atmosphere. At 128 km Marmo *et al.*[16] also determined the diffusion coefficient for sodium from optical observations. The value obtained was 7×10^{7} cm^2/sec which is not consistent with that found for caesium from the radio backscatter experiment, since the coefficient should be greater for sodium than for caesium.*

The final diffusion coefficients shown in Fig. 4 are those deduced from Zimmerman and Champion's[17] results for caesium atoms by allowing for the mass difference of sodium and caesium atoms. Again, the diffusion coefficients found for sodium may be in error by ~20 per cent. Even so, they follow the general pattern of the other results.

The curve AA′ in Fig. 4 has been drawn

* It has been pointed out by Marmo (private communication) that the value of $2{\cdot}5\times10^{8}$ cm^2/sec found for caesium from the radio backscatter observations is expected to be more reliable than that of 7×10^{7} cm^2/sec obtained for sodium from the optical observations. This is mainly due to the limited calibration techniques available in the optical experiment[16].

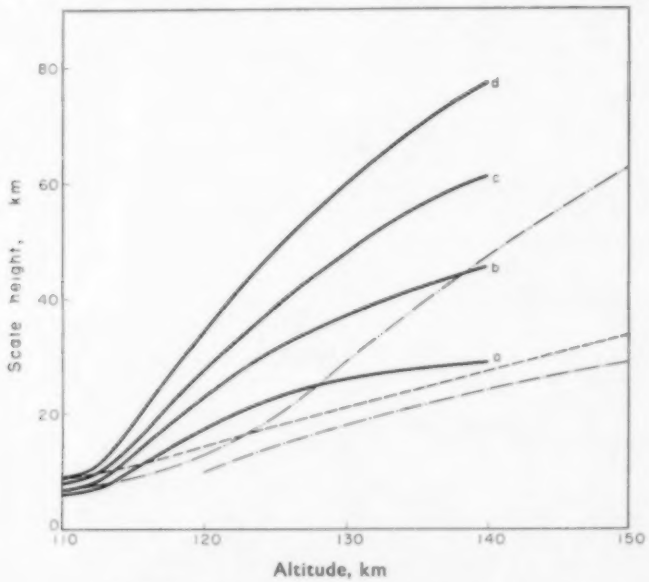


Fig. 5. Variation with altitude of atmospheric scale heights.
———— Scale heights calculated from "Skylark 16" data.
Curve a with H at 110km = 6 km
Curve b " " " " = 7 km
Curve c " " " " = 8 km
Curve d " " " " = 9 km
— — — Minzner et al.[12]
— . — . —Horowitz and La Gow[19]
— . . — . . —La Gow et al.[18]

through the results of the present investigation. It is interesting to determine from this curve values of the atmospheric scale height, H. If equation (9) is written as

$$D = D_0 \sqrt{T}/N\sqrt{\mu} \tag{11}$$

and if H has a constant gradient of $dH/dz = \beta$ over a small range of altitudes, then

$$D_2/D_1 = \{(H_1 + \beta\delta z)/H_1\}^{[(2+3\beta)/2\beta]} \tag{12}$$

where D_1 and H_1 are the values of the diffusion coefficient and scale height at an altitude z and D_2 and $H_1 + \beta\delta z$ are the corresponding values at an altitude $z + \delta z$. The variation of M over the range of δz has been assumed small, and the change with z of g, the acceleration due to gravity, has been neglected. Using values of D taken from curve AA', values of H and β have been calculated from equation (12) taking $H = 6$, 7, 8 and 9 km at an altitude of 110 km. The resulting variations of H with altitude are shown in Fig. 5 (curves *a*, *b*, *c*, *d*) together with that used in the 1959 A.R.D.C. Standard Atmosphere and with the results of two rocket flights at Churchill[18] and one at White Sands[19]. It is seen that the value chosen for the scale height at 110 km has a marked effect on the values found for higher altitudes, although the general shape of the curves is similar. From consideration of temperatures deduced from the curves *a*, *b*, *c* and *d* at altitudes ~ 140 km, it seems reasonable to assume that the scale height at 110 km was between 7 and 8 km. For these two values the corresponding temperatures at altitudes up to 140 km are given in Table 1 for a constant atmospheric mean molecular mass of 28·97. The scale height gradients for $H_{110} = 7$ and 8 km are given in Table 2 which reveals a marked increase in β at about 112 km, β increasing to a maximum of between 2 and 2·5 km/km at about 118 km. Above this altitude β begins to reduce. The large values of β between 112 and 130 km are indicative of the increase of temperature which occurs in this region. The scale height gradients between 112 and 130 km are between three and four times greater than those of the 1959 A.R.D.C. Standard Atmosphere, but are comparable with the value of about 1·8 km/km deduced from the scale heights of La Gow *et al.* for the firing of 29 July, 1957[18]. The marked increase in the scale height gradient, observed by La Gow *et al.*, began near 120 km and reached its maximum at about 130 to 135 km, i.e. approximately 15 km higher than in the present case. This difference of 15 km may represent the difference in the altitudes of maximum temperature.

From the scale heights given in curves *b* and *c* of Fig. 5 atmospheric densities at altitudes between 110 and 140 km have been calculated from

$$\rho_2/\rho_1 = (H_2/H_1)^{-(1+\beta)/\beta} \tag{13}$$

taking $\rho_1 = 6 \times 10^{-5}$ g/m^3 at 110 km. The

Table 1. Calculated temperatures for altitudes between 110 and 140 km.

Altitude (km)	110	112	114	116	120	125	130	135	140
Temperature (°K) $H_{110} = 7$ km	230	260	365	485	750	1025	1220	1390	1490 to 1540
Temperature $H_{110} = 8$ km	265	310	440	570	905	1275	1580	1830	1990 to 2050

Table 2. Scale height gradients for altitudes between 110 and 140 km.

Altitude (km)	111·0	113·0	115·0	118·0	122·5	127·5	132·5	137·5
β (for $H_{110} = 7$ km)	0·45	1·6	1·8	2·0	1·65	1·2	1·0	0·6 to 0·75
β (for $H_{110} = 8$ km)	0·65	1·8	2·0	2·5	2·25	1·85	1·5	1·0 to 1·2

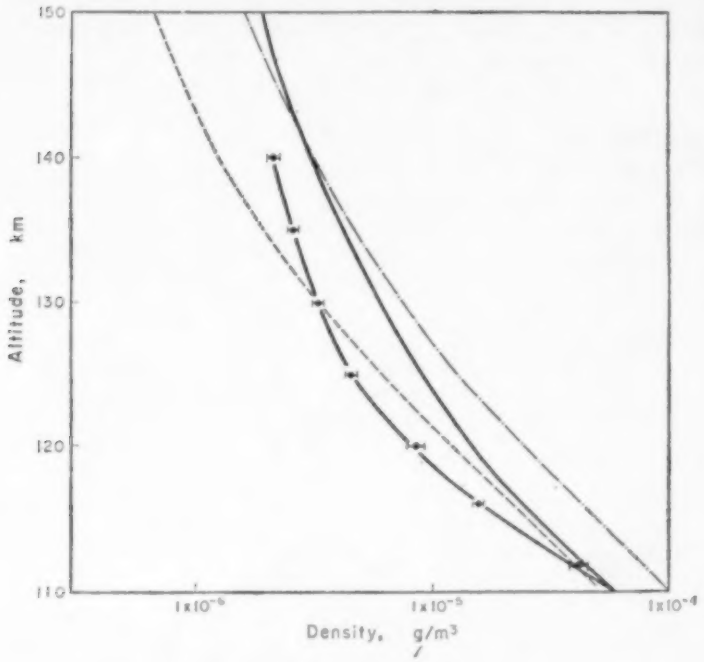


Fig. 6. Variation with altitude of atmospheric mass density.
|—●—| "Skylark 16" data
———— ARDC model Atmosphere[12]
— . — . — Michnevitch[20]
— — — Horowitz and La Gow[19]

densities so calculated differed by less than 10 per cent for a particular altitude and the mean values are shown in Fig. 6 together with the 1959 A.R.D.C. curve and the results of two other rocket determinations[19,20]. It is seen that the densities found from the present work are in general agreement with those found from other determinations but are rather lower. Between 110 and 125 km the rate of decrease in density is greater than for the A.R.D.C. curve while above 125 km the opposite is true.

5. CONCLUSIONS

A large part of the differences between the diffusion coefficients found in the present investigation and those found by others using similar methods is probably due to the simplifying assumptions made and to the experimental errors involved. Further, the difference, by as much as a factor of three, between the present diffusion coefficients and those calculated from the 1959 A.R.D.C. Model Atmosphere may be partly due to the uncertainty in the value of the diffusion cross-section for sodium used in equation (9). However, the comparison of the diffusion coefficients, scale heights and mass densities found from the sodium trail laid by Skylark 16 with other data does indicate variations in atmospheric structure. It is concluded that future experiments, in which attempts might be made to allow for the effects of winds, self-absorption and multiple scattering, would be an effective method of studying the variations with altitude of the atmospheric scale height and mass density. The altitude limits for such experiments are somewhat uncertain; below 110 km the diffusion is due increasingly to eddy diffusion (Blamont and de Jager[9]) while at great altitudes difficulties arise from the rarity of the atmosphere. Shklovskii and Kurt[14] have ap-

plied the method at an altitude of 430 km and deduced a value of the atmospheric density in agreement with those found from satellite observations.

It is intended that the present sodium trail experiments will be continued at Woomera, the diffusion coefficient measurements being extended to altitudes of 200 km.

Acknowledgments—The author wishes to acknowledge his gratitude to Professor D. R. Bates, F.R.S., in whose Department the work was carried out, for his great interest and encouragement, to Drs. A. Dalgarno and A. H. Jarrett for valuable discussions, and to Mr. G. T. Best for his co-operation. The work would not have been possible without the assistance of members of the Weapons Research Establishment, particularly of Messrs. J. S. Colville and B. Rudd, and of members of the Royal Aircraft Establishment, Farnborough.

REFERENCES

1. D. R. Bates, *J. Geophys. Res.* **55,** 347 (1950).
2. G. V. Groves, *Nature, Lond.* **187,** 1001 (1960). *Space Research*, p. 144 (Ed. H. K. Kallam Bijl) North Holland, Amsterdam (1960).
3. J. A. Rees, *Space Research*, p. 207 (Ed. H. K. Kallman Bijl) North Holland, Amsterdam (1960).
4. J. A. Rees and D. Parkinson (in preparation).
5. A. Dalgarno (private communication).
6. J. Blamont, C. R. *Acad. Sci. Paris*, **14,** 1248 (1959). *Space Research*, p. 199 (Ed. H. K. Kallam Bijl) North Holland, Amsterdam (1960).
7. E. Manring, J. F. Bedinger and H. B. Pettitt, *J. Geophys. Res.* **64,** 587 (1959).
8. J. Blamont and C. de Jager, *Ann. Géophys.*, to be published.
9. E. Manring, *G.R.D. Report on Project 'Firefly* 1960' **1,** 2034 (1960).
10. J. Blamont (private communication).
11. G. C. A. results quoted by S. P. Zimmerman and K. S. W. Champion. G.R.D. *Report on Project 'Firefly* 1960' **1,** 2029 (1960).
12. R. A. Minzner, K. W. Champion and H. Pond, *The* 1959 *ARDC Model Atmosphere*, U. S. Air Force Surveys in Geophysics, No. 115, ARCRC (1959).
13. S. Chapman and T. G. Cowling, *Mathematical Theory of Non-Uniform Gases* p. 245, Cambridge University Press, Cambridge (1939).
14. I. S. Shklovskii and V. G. Kurt, *Iskusstvennye Sputniki Zemli*, **3,** 66 (1959).
15. F. F. Marmo, L. M. Aschenbrand and J. Pressman, *J. Amer. Rocket Soc.* **30,** 523 (1960).
16. F. F. Marmo, J. Pressman, E. R. Manring and L. Aschenbrand, *Planet. Space Sci.* **2,** 174 (1960).
17. S. P. Zimmerman and K. S. W. Champion, *G.R.D. Report on Project 'Firefly* 1960' **1,** 2025 (1960).
18. H. E. La Gow, R. Horowitz and J. Ainsworth, *Space Research*, 164 (Ed. H. K. Kallman Bijl) North Holland, Amsterdam (1960).
19. R. Horowitz and H. E. La Gow, *J. Geophys. Res.* **62,** 57 (1957): **63,** 757 (1958).
20. V. V. Michnevitch, Paper to 5th Assembly CSAGI, Moscow (1958).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 43–48. Printed in Great Britain.

DETERMINATION OF THE EARTH'S GRAVITATIONAL POTENTIAL FROM SATELLITE ORBITS

D. E. SMITH

D.S.I.R. Radio Research Station, Slough

(*Received* 13 *May* 1961)

Abstract—From the motions of the orbital planes of four satellites, values for the second, fourth and sixth harmonics in the earth's gravitational potential have been obtained. Allowance is made for atmospheric, lunar and solar perturbations and the results are compared with those of other authors.

1. INTRODUCTION

The principal effects of the earth's oblateness on the orbit of a near earth satellite are the rotation of the orbital plane and of the major axis within the orbital plane. From observations of the satellite and the determination of the elements of the orbit over a long period, the actual oblateness perturbations can be deduced to quite high accuracy and, under the assumption that the shape of the earth closely resembles an equipotential surface of an oblate spheroid, it is possible to evaluate the terms in the expansion of the earth's potential function. Estimates of the even harmonics have been made by several authors[1–8] as a result of observations of these perturbations. In addition to these principal effects, the inclination of the orbital plane and the eccentricity are known to undergo oscillations of period similar to that of the rotation of the major axis, and these effects have led to estimates of the odd harmonics in the potential function[1, 5, 9]. The potential function contains an infinite number of terms and for it to be possible to estimate the values of a few of these it is necessary to assume those of higher order to the zero. For this reason, with the exception of King-Hele[7] all authors so far have neglected terms above the fourth. By determining the rate of change of right ascension of the ascending node of three satellites, he has obtained values for the second, fourth and sixth harmonics, neglecting the eighth harmonic and above.

In this paper, values of the second, fourth and sixth harmonics are derived from the rates of rotation of the orbital planes of four satellites other than those used by King-Hele[10], in an attempt to obtain further independent values for the J_n under the same conditions.

2. METHOD OF ANALYSIS

The potential function of the earth at a point outside it is a solution of Laplace's equation

$$\nabla^2 U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0 \qquad (1)$$

where U is the potential function.

If we assume that the earth's potential is independent of longitude and, further, that the equatorial plane passes through the centre of mass of the earth, then the most convenient and explicit solution of Laplace's equation at an exterior point distance r from the centre of the earth and with co-latitude θ, is

$$U = \frac{GM}{R}\left\{\frac{R}{r} - \sum_{n=2}^{\infty} J_n \left(\frac{R}{r}\right)^{n+1} P_n(\cos\theta)\right\} \qquad (2)$$

where G is the gravitational constant
M is the mass of the earth
R is the earth's equatorial radius
$P_n(\cos\theta)$ is the Legendre polynomial of order n defined as

$$P_n(Z) = \frac{1}{2^n} \sum_{0}^{[n/2]} \frac{(-1)^r (2n-2r)!}{r!\,(n-2r)!\,(n-r)!} Z^{n-2r}$$

and J_n are constants to be evaluated.

From this expression for U, Merson[11] and King-Hele[12] have obtained expressions for the rate of rotation of the orbital plane of a near earth satellite. Combining these expressions we obtain

$$\frac{d\Omega}{dt} = -\frac{3}{2}\left(\frac{R}{\bar{r}}\right)^2\left(\frac{GM}{a^3}\right)^{\frac{1}{2}}\cos\alpha\left\{\frac{3}{2}\left(\frac{R}{\bar{r}}\right)^2\left(\frac{19}{12}\sin^2\alpha - 1\right)J_2^2 + J_2 + \frac{1}{4}\left(\frac{R}{\bar{r}}\right)\frac{e\sin\omega}{\sin\alpha}(15\sin^2\alpha - 4)J_3\right.$$

$$-\frac{5}{2}\left(\frac{R}{\bar{r}}\right)^2\left[\left(1-\frac{7}{4}\sin^2\alpha\right)\left(1+\frac{3}{2}e^2\right)+\left(\frac{7}{4}\sin^2\alpha - \frac{3}{4}\right)e^2\cos 2\omega\right]J_4$$

$$+\frac{5}{2}\left(\frac{R}{\bar{r}}\right)^3\frac{e}{\sin\alpha}\left[\left(1-\frac{21}{2}\sin^2\alpha+\frac{105}{8}\sin^4\alpha\right)(1+\tfrac{3}{4}e^2)\sin\omega+\tfrac{7}{8}\left(1-\frac{15}{8}\sin^2\alpha\right)e^2\sin 3\omega\right]J_5$$

$$+\frac{35}{8}\left(\frac{R}{\bar{r}}\right)^4\left[\left(1-\frac{9}{2}\sin^2\alpha+\frac{33}{8}\sin^4\alpha\right)\left(1+5e^2+\frac{15}{8}e^4\right)\right.$$

$$\left.-\frac{5}{2}\left(1-6\sin^2\alpha+\frac{99}{16}\sin^4\alpha\right)(1+\tfrac{1}{2}e^2)e^2\cos 2\omega-\frac{15}{32}\left(1-\frac{33}{20}\sin^2\alpha\right)e^4\sin^2\alpha\cos 4\omega\right]J_6$$

$$\left.-\frac{105}{16}\left(\frac{R}{\bar{r}}\right)^6\left[1-\frac{33}{4}\sin^2\alpha+\frac{143}{8}\sin^4\alpha-\frac{715}{64}\sin^6\alpha\right]J_6+0\,(J_2^2e^6,\ J_2J_3,\ J_2J_4,\ldots)\right\} \qquad (3)$$

where $\bar{r}=a(1-e^2)$ and α is the geocentric inclination of the orbit to the equator, a is the semi-major axis, e the eccentricity and ω the argument of perigee.

The determination of the potential function involves evaluating the J_n constants and this can be done with the aid of equation (3). The coefficients of the J_n terms are evaluated for several satellites and the right hand side of equation (3) equated to the observed values of $d\Omega/dt$. In this way, each determination of $d\Omega/dt$ can provide a relationship between the J_n terms.

If reasonably accurate observed values of the right ascension of the ascending node are available over a period of two, three or four months, then it is possible to estimate a mean value for $d\Omega/dt$ on a mean date to high accuracy. Because some of the orbital elements exhibit long and short period oscillations, the orbital elements used in equation (3) are not the osculating elements but specially defined "smoothed elements". The shorter oscillations have a period of one revolution whilst the longer oscillations have a period similar to that of the rotation of the major axis of the orbit. The short period oscillations are discussed in detail in reference[13]. A practical example of the long period oscillation is the variation of the inclination of the orbit to the equator. In the case of Vanguard II satellite, whose orbital inclination is about 32 degrees, the variation has an amplitude of about 0·017 degrees and a period of about 85 days. If a mean value of $d\Omega/dt$ is determined over a period of a few months then the value for the inclination on the mean date used in this analysis is not the value of the osculating element but a mean inclination averaged over the whole period under consideration.

Besides the right ascension of the ascending node and the inclination, the only other orbital element that is required to great accuracy is the semi-major axis, a. This is obtained to an accuracy of about 0·003 per cent directly from the anomalistic period. In the case of satellites that are observed regularly from the United Kingdom or Western Europe, graphs of the variation of the nodal period with time are kept at the Radio Research Station and the value of the nodal period at any given instant can be obtained to at least 0·005 of a minute. From the nodal period, a simple calculation gives the anomalistic period and hence the semi-major axis. An error of $\pm 0{\cdot}005$ minutes in the nodal period corresponds to an error of about $\pm 0{\cdot}25$ km in the semi-major axis.

3. DATA USED

Of the ten satellite orbits analyzed, only those of Sputnik III satellite, Transit 1B rocket, Tiros 1 satellite and Vanguard II were known

to sufficient accuracy for them to be used in the evaluation of the J_n constants. The relatively short lifetime of the rocket of Sputnik III prevented data from this source being used since the changes in the orbital elements over a given period were much larger for the rocket than the satellite. In particular, the changes in the elements of the orbit of the rocket could not be assumed linear and this caused difficulty when trying to determine the mean orbital elements on the mean date.

The reduction of a large number of visual observations on Sputnik III satellite made in the United Kingdom and Western Europe together with orbital elements obtained by Cornford, King-Hele, Merson[14] provided a series of values for the elements extending over a large proportion of the lifetime of the satellite. Because the inclination of the orbit of Sputnik III was about 65°, the perigee position rotated very slowly, taking about 750 days for one revolution and consequently it was impossible to discern the long period oscillations of the elements (the lifetime of Sputnik III satellite was 692 days). The right ascension of the ascending node was plotted against time and a value for the rate of change obtained for 26th August 1958 by subtracting two values of the right ascension on opposite sides of this date and dividing by the time interval.

Until the launch of the Echo 1 satellite, Transit 1B rocket had been one of the most frequently observed object from the British Isles and Western Europe. Theodolite observations on Transit 1B rocket from Meudon Observatory in France and Jokioinen in Finland together with ordinary visual observations from the United Kingdom enabled values of the right ascension of the ascending node, nodal period and inclination to be obtained during the months of May, June and July 1960. The values of the eccentricity and argument of perigee were obtained from the elements derived by the American Prediction Service, Spacetrack.

Tiros 1 satellite has also been observed on several occasions from the British Isles but these observations were too few in number to provide independent values for the right ascension of the ascending node. For this reason all the orbital elements were obtained from the Spacetrack Bulletin Predictions. The values for the elements were plotted against time but no long-period variations were discernable. A set of "mean" elements, together with the rate of rotation of the orbital plane, were obtained in this way for 3rd June 1960.

The only other satellite for which it was possible to determine the "mean" orbital elements and the rate of rotation of the orbital plane to sufficient accuracy was Vanguard II. Because the satellite is not transmitting and has only been observed visually from the United Kingdom on a few isolated occasions, it was necessary to rely on the elements computed by the Smithsonian Astrophysical Observatory for their predictions. Curves of the variation of the right ascension of the ascending node, inclination, argument of perigee and eccentricity with time were plotted and provided a value of $d\Omega/dt$ on 29th March 1960; with the appropriate "mean" elements. As explained earlier, the inclination appeared to oscillate with a period of about 85 days and the mean value differed from the osculating value on 29th March 1960 by about 0·015°. Similarly, the eccentricity was seen to oscillate but with a period of about 65 days and amplitude of about 0·0004.

In Table 1, the mean orbital elements are given for each of the four satellites together with the corresponding observed rate of rotation of the ascending node. The satellites are tabulated in order of increasing inclination.

These observed values of the rates of rotation of the orbital planes given in Table 1 include the rotational effects produced by the perturbing forces of the sun, moon and the earth's atmosphere. Although these effects are small compared with those of oblateness it is necessary to make allowance for them. Kozai[14] has derived an expression for the effect of the sun and moon and Cook[16] for the atmosphere on the rotation of the nodes. Table 2 shows the computed magnitudes of the combined effects of the sun and moon on the orbits of the four objects considered here. The atmospheric effects were found to be about 200 times less than the luni-solar contribution.

Table 1. Mean orbital elements

Satellite	Designation	Date	Mean orbital elements				Rate of rotation of ascending note $\left(\frac{d\Omega}{dt}\right)$ (degs/day)
			Semi-major axis (km)	Eccentricity	Orbital inclination (deg)	Argument of perigee (deg)	
Vanguard II satellite	1959 Alpha 1	29 March 1960	8307·17	0·16485	32·880	+112·3	−3·5114
Tiros I satellite	1960 Beta 2	3 June 1960	7100·26	0·00448	48·410	+ 16·0	−4·5443
Transit 1B rocket	1960 Gamma 1	8 June 1960	6889·76	0·02810	51·292	+108·0	−4·7634
Sputnik III satellite	1958 Delta 2	26 August 1958	7355·66	0·10401	65·170	+ 21·2	−2·5956

Table 2.

Satellite	Combined rotation of orbital plane by moon and sun (degs/day)
Vanguard II satellite	−0·00039
Tiros I satellite	−0·00023
Transit 1B rocket	−0·00021
Sputnik III satellite	−0·00016

With the aid of Table 1 it is possible to evaluate the coefficients of the J_n terms in equation (3) thereby obtaining an equation connecting the J_n for each of the four satellites. Thus it is possible to obtain four equations connecting six unknowns. If values are assumed for J_3 and J_5 then it should be possible to determine values for J_2, J_4, J_6 and J_8. However, this still leaves the J_2^2 term, but this may also be allowed for by assuming a value for J_2. Because J_2 itself is small, even an approximate value of J_2 will give J_2^2 to sufficient accuracy.

Several authors have estimated values for J_3 and J_5 and those used here are by O'Keefe, Eckels and Squires[1] who obtained

$$10^6 J_3 = -2{\cdot}4 \pm 0{\cdot}3$$
$$10^6 J_5 = -0{\cdot}1 \pm 0{\cdot}1$$

The existence of these odd harmonics, J_3, J_5 etc., in the earth's gravitational field shows that the earth is, in fact, pear-shaped. Had all these odd order harmonics been zero, symmetry would have existed between the northern and southern hemispheres.

If J_2 is assumed to be $(1083 \pm 2) \times 10^{-6}$, which includes the complete range of values so far obtained for J_2, then $J_2^2 = (1{\cdot}173 \pm 0{\cdot}004) \times 10^{-6}$ which is much more accurate than J_3 or J_5.

4. RESULTS

The substitutions were carried out as described for J_2^2, J_3 and J_5 in equation (3) and four equations with four unknowns obtained, J_2, J_4, J_6 and J_8; the equations being most easily expressed in matrix form as

$$\begin{matrix}\text{Vanguard II}\\ \text{Tiros I}\\ \text{Transit 1B rocket}\\ \text{Sputnik III}\end{matrix}\begin{bmatrix}1 & -0{\cdot}7920 & +0{\cdot}0426 & +0{\cdot}2614\\ 1 & -0{\cdot}0426 & -0{\cdot}6453 & +0{\cdot}0788\\ 1 & +0{\cdot}1414 & -0{\cdot}6835 & -0{\cdot}3396\\ 1 & +0{\cdot}8511 & +0{\cdot}2373 & -0{\cdot}2644\end{bmatrix}\begin{bmatrix}J_2\\ J_4\\ J_6\\ J_8\end{bmatrix} = \begin{bmatrix}1{\cdot}08428 \times 10^{-3}\\ 1{\cdot}08270 \times 10^{-3}\\ 1{\cdot}08257 \times 10^{-3}\\ 1{\cdot}08213 \times 10^{-3}\end{bmatrix}$$

From these equations it should have been possible to obtain values for each of the J_n but unfortunately the terms on the right hand side of the equation for Tiros satellite and Transit rocket are nearly equal. Consequently, it was only possible to use Tiros and Transit separately and to form two 3×3 matrix equations under the assumption that the J_8 term was zero.

These two matrix equations gave

$10^6 J_2 = 1083{\cdot}16 \pm 0{\cdot}20 \quad 10^6 J_2 = 1083{\cdot}14 \pm 0{\cdot}20$
$10^6 J_4 = -1{\cdot}4 \pm 0{\cdot}3 \quad 10^6 J_4 = -1{\cdot}4 \pm 0{\cdot}3$
$10^6 J_6 = +0{\cdot}6 \pm 0{\cdot}6 \quad 10^6 J_6 = +0{\cdot}8 \pm 0{\cdot}6$

with mean values of

$$10^6 J_2 = 1083{\cdot}15 \pm 0{\cdot}20$$
$$10^6 J_4 = -1{\cdot}4 \pm 0{\cdot}3$$
$$10^6 J_6 = +0{\cdot}7 \pm 0{\cdot}6$$

In these determinations, the errors quoted are those from observational errors and do not include any errors that may occur by neglecting any of the higher J_n terms.

Table 3 lists the results obtained by several other investigators for the even harmonics in the gravitational field. The author's results are quoted at the bottom of the table together with the mean values obtained from five 2×2 matrix equations when both J_6 and J_8 were assumed zero.

5. CONCLUSIONS

From Table 3 it is evident that the results described in this paper agree reasonably well with those obtained by King-Hele. The J_2 value perhaps appears a little high but the J_4 and J_6 terms agree well. It is interesting to note the reasonably large variations in the values of J_4 obtained by other authors in which J_6 has been assumed zero. In an attempt to discover

Table 3. Computed Values of J_2, J_4, J_6

Author	Satellites used	J_2 in units of 10^{-6}	J_4 in units of 10^{-6}	J_6 in units of 10^{-6}
King-Hele[7, 10]	1958 β_2 1957 β 1959 ι_1	1082·79 (±0·15)	−1·4 (±0·2)	+0·9 (±0·8)
King-Hele[4]	1958 β_2 1957 β 1958 ϵ	1083·1 (±0·2)	−1·4 (±0·2)	+0·1 (±1·5)
O'Keefe, *et al.*[1]	1958 β_2	1082·5 (±0·1)	−1·7 (±0·1)	— —
Kozai*	1958 β_2 1959 η 1959 ι_1	1082·19 (±0·03)	−2·13 (±0·05)	—
Buchar*	1957 β 1958 δ_2 1958 α 1958 β_2 1959 α_1	1083·0 (±0·4)	−1·1 (±0·5)	—
Zhongolovich[8]	1957 β_2 1958 δ_1 1958 δ_2	1083·3 (±0·7)	−4·1 (±0·7)	—
Jacchia[3]	1957 β 1958 β_2	1082·7 (±0·7)	−2·1 (±0·5)	—
Lecar, *et al.*[2]	1958 β_2	1082·1 (±0·3)	—	—
Smith	1958 δ_2 1959 α_1 1960 β_2 1960 γ_1	1083·15 (±0·2) 1082·8 (±0·2)	−1·4 (±0·3) −1·3 (±0·3)	+0·7 (±0·6) —

* Papers circulated at the General Assembly of the International Union of Geodesy and Geophysics, Helsinki, July/August 1960.

the effect of J_6 on J_2 and J_4, the author has included at the bottom of Table 3 results obtained from the matrix equation when both J_6 and J_8 are assumed zero. The mean value of J_2 has been lowered while that of J_4 has hardly changed; but perhaps what was more significant was the scatter of the values for J_4 which were obtained; these lay between $-0{\cdot}6\times10^{-6}$ and $-2{\cdot}1\times10^{-6}$. These results may be an indication that the effect of J_6 on J_4 and J_2 is quite appreciable and poses the question as to the effects of J_8 and J_{10} etc.

Acknowledgements—The author would like to take this opportunity of acknowledging the co-operation of the observers whose reports made a valuable contribution to the data used for this work.

The work described here was carried out as part of the programme of the Radio Research Board and is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

REFERENCES

1. J. A. O'Keefe, A. Eckels and R. K. Squires. *Astr. J.* **64,** 245 (1959).
2. M. Lecar, J. Sorenson and A. Eckels. *J. Geophys. Res.* **64,** 209 (1959).
3. L. G. Jacchia. *Smithsonian Astrophysical Observatory. Special Report No. 19,* 1 (1958).
4. D. G. King-Hele and R. H. Merson. *Nature Lond.* **183,** 881 (1959).
5. Y. Kozai. *Smithsonian Astrophysical Observatory, Special Report No. 22,* 1 (1959).
6. R. H. Merson and D. G. King-Hele. *Nature, Lond.* **182,** 640 (1958).
7. D. G. King-Hele. *Nature, Lond.* **187,** 490 (1960).
8. I. D. Zhongolovich. *Bulletin of Optical Observing Stations, Astronomical Council of the Academy of Sciences, U.S.S.R.,* **2,** 1 (1960).
9. J. O'Keefe, A. Eckels and R. K. Squires. *Science,* **129,** 565 (1959).
10. D. G. King-Hele. *Geophys. J.,* **4,** 3 (1961).
11. R. H. Merson *Geophys. J.,* **4,** 17 (1961).
12. D. G. King-Hele. *Proc. Roy. Soc.* A, **247,** 49 (1958).
13. D. G. King-Hele and R. H. Merson. *J. Brit. Interplan. Soc.,* **16,** 446 (1958).
14. E. C. Cornford, D. G. King-Hele and R. H. Merson. *Proceedings of the 7th Anglo-American Aeronautical Conference,* New York, 1959, Pergamon Press.
15. Y. Kozai. *Smithsonian Astrophysical Observatory Special Report No. 22,* 7 (1959).
16. G. E. Cook. *Proc. Roy. Soc.* A, **261,** 246 (1961).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 49–58. Printed in Great Britain.

THE AURORA OF 27 NOVEMBER 1959 AT COLLEGE, ALASKA INCLUDING OBSERVATIONS OF A HIGH ALTITUDE RED ARC

M. H. REES
Department of Applied Mathematics, The Queen's University, Belfast*

and

C. S. DEEHR
Physics Department, University of Washington, Seattle†

(*Received* 28 *June* 1961)

Abstract—Spectrophotometric observations on a rare high altitude red auroral arc were obtained at College, Alaska, in the evening of 27 November 1959. Photometric and photographic triangulation yielded a height of about 350 km for the region of maximum luminosity. The arc can be characterized spectroscopically by the OI λλ6300-6364 lines and the N_2^+ 1 NG bands, other usual auroral emission features being notably absent. Though the hydrogen lines are present, their identification with the red arc is not certain. The maximum intensity of OI λ6300, 66·7 kR,†† was reached when the arc passed through the magnetic zenith. The arc was approximately aligned with the geomagnetic field and shifted approximately 10° in latitude, north to south, in a time interval of one hour. An intense and active aurora followed the appearance of the red arc and lasted throughout the night, but did not exhibit any unusual features.

INTRODUCTION

The high latitude red arc is a rare phenomenon. Only visual and photographic observations of a few such arcs have been reported[1, 2], while the spectral details have, up to the present, been unknown.

On 27 November 1959 at 1642 AST (Alaska standard time, 150° WMT) the all-sky camera at Point Barrow, Alaska, recorded a faint auroral arc at 30° zenith distance to the geomagnetic north. At College, Alaska, a photometer which had just completed measurements of the enhancement of the sodium *D* lines in the twilight detected the arc during an almucantar sweep (in λ6300) at about 1645 AST. A programme of spectroscopic, photometric and all-sky camera observations was in effect by 1700 AST and the arc was recorded as it passed into an intense predominantly green aurora which formed in the southern sky at College and moved further south. The extensive experimental apparatus for auroral research built up for the IGY programme at College was employed in this investigation[3].

EXPERIMENTAL

The photometric data were taken with a birefringent filter photometer basically designed by Dunn and Manring[4] and developed at the National Bureau of Standards for the IGY airglow and aurora programme. It has been described by Megill[5]. The high rejection of continuous light by the birefringent filter required the use of an auxiliary photometer for absolute calibration. The latter used a narrow pass band interference filter. Our calibration was compared with that of Roach[6] who carried a portable calibrating photometer to several airglow-aurora observatories during the IGY. From a variety of arguments we believe that the numbers should be correct to better than a factor of two.

The photometer ordinarily was programmed to sweep continuously in the geomagnetic N–S meridian providing data on the luminosity as a function of zenith distance. While the red

* *Present address*: Department of Physics, University of Colorado, Boulder, Colorado.

† *Present address*: Geophysical Institute, University of Alaska, College, Alaska.

†† kR = kilorayleigh.

arc was in the zenith the telescope was pointed in that direction and rotated about its axis. The built-in fixed polaroid allows using the instrument in this manner as a detector of polarization. It was also pointed to the zenith for calibration runs. Data were recorded continuously on strip chart paper. The intensities were corrected for extinction by the lower atmosphere; details are given by Deehr[7].

The meridian mirror auroral spectrograph[8] was operated throughout the night taking 15-min exposures in the visible region of the spectrum of a narrow strip of sky in the geomagnetic meridian. The sensitivity of the instrument is such that the routine 15-min exposures are sufficient to record the chief auroral emission features, while the relatively good time resolution shows the motion of auroral forms in the geomagnetic meridian, revealing their spectral characteristics. The spectra were calibrated against the photometric records at λ6300, while a relative calibration at other wavelengths allowed deduction of the absolute intensity of several important spectral features. Calibration spectra were exposed for the same length of time as the auroral spectra to avoid introducing errors due to the failure of the reciprocity law. The calibration curve obtained for long exposures is indeed considerably different from the usual curve derived from short exposures of less than 15 sec. We believe that this is the principal cause of the uncertainty usually attached to the intensities derived from photographic spectra. Details of the calibration procedure and the corrections applied for extinction and scattering are described by Deehr[7, 9].

All-sky cameras[10] were operated throughout the night at College and Point Barrow, one 15-sec exposure being obtained at the beginning of each minute.

THE RED ARC

The all-sky camera pictures provided a record of the motions and overall morphology of the aurora, beginning with the detection of the faint red arc approximately 300 km north of Point Barrow at 1642 AST. By the time the red arc had reached the zenith at College (approximately 1710 AST) it had acquired red ray structure which showed east to west motion in the arc. The human eye is relatively insensitive to λ6300 and the red colour does not normally become visible until the intensity reaches about 30 kR*. At 1645 AST the arc was seen in the north to be distinctly red and when photometric observations were begun a few minutes later the recorded intensity was 25 kR.

The slant intensities in the geomagnetic meridian of λ6300 and λ5577 during the period in which the red arc was visible from College are given in Table 1. The morphology of the photometric traces, after calibration and corrections were applied, is shown in Fig. 1. The level of intensity rises slowly with elevation suggesting a diffuse rather than a sharp lower border, or conceivably the presence of another arc farther north. The decrease in intensity at the top of the arc seems to be more abrupt. The peak photon intensity and the zenith distance of the peak are shown in Fig. 2 for each meridian sweep. It appears that the arc travelled much more rapidly before reaching the magnetic zenith than afterwards. The peak intensity occurred when the arc was close to the magnetic zenith indicating that the arc was field aligned and that it extended over a considerable range of height. Lack of meridional data during the time the arc was in the geographic zenith is due to an attempt to measure polarization of λ6300 in the arc by rotating the photometer about its optic axis. No measurable polarization could be found. Duncan[11] has reported one case of polarization of a middle latitude subvisual λ6300 arc. It should be noted that the photometer could monitor only one wavelength at a time, thus the records of λ6300 and λ5577 were taken one minute apart. Nevertheless, one of the striking features evident in Table 1 and Fig. 1 is the great enhancement of the red lines over the green line in the arc. However, this was not the case for another arc which formed in the southern sky at about 1723 AST, as shown in the figure.

Auroral arcs in Alaska are generally aligned

* $1\,kR = 10^9$ photons $cm^{-2}\,sec^{-1}$ $(column)^{-1}$.

Table 1. Slant Intensity in Kilorayleighs—College, Alaska, 27 November, 1959

Wave-length	Time 150 WMT h m s	Zenith Distance																	Maxima	
		80°N	70°N	60°N	50°N	40°N	30°N	20°N	10°N	0	10°S	20°S	30°S	40°S	50°S	60°S	70°S	80°S	ZD	kR
λ6300	165300	5·4	5·1	14·9	23·0	11·2	3·4	—	—	—	—	—	—	—	—	—	—	—	50°N	
	1655	5·4	7·1	16·8	19·5	6·9	3·4	2·5	2·5	2·5	—	—	—	—	—	—	—	—	50°N	
	1656	4·0	4·1	12·1	20·4	11·2	4·3	1·7	1·7	0·83	—	—	—	—	—	0·93	2·0	4·1	50°N	
	1657	4·1	5·1	8·4	13·3	14·7	6·8	2·5	1·7	0·83	0·83	0·85	0·85	0·85	0·89	0·93	1·02	2·7	43°N	15·7
	1658	4·1	3·1	5·6	10·6	16·4	12·0	4·2	2·5	1·7	0·83	0·85	0·85	0·85	0·89	0·93	2·0	4·1	40°N	
	1659	4·3	4·9	6·7	9·2	15·9	17·1	8·5	2·3	1·7	1·3	1·4	1·4	1·4	1·4	1·4	1·5	1·6	35°N	17·9
	1702	6·0	5·3	6·0	9·2	19·0	28·8	*	15·0	2·7	1·3	1·4	1·7	1·7	1·1	1·9	2·4	3·8	20°N	
	1705	5·4	4·1	2·8	3·5	4·3	7·7	17·8	23·3	15·8	1·7	0·85	0·85	0·85	0·89	0·93	1·0	2·7	10°N	
	1706	5·4	4·1	3·7	4·4	5·2	7·7	14·4	25·0	23·3	5·0	1·7	0·85	0·85	0·89	0·93	1·0	2·7	5°N	28·3
	1707	4·1	3·1	2·8	2·7	3·4	7·7	15·3	25·0	15·0	2·5	1·7	1·7	1·7	0·89	0·93	2·0	4·0	10°N	
	1708	2·7	2·0	1·9	2·7	4·3	8·6	13·6	20·8	19·2	4·2	1·7	0·85	0·85	0·89	0·93	1·0	1·4	5°N	24·1
	1709	2·7	2·0	1·9	1·8	2·6	5·1	12·7	19·2	17·5	1·7	0·85	0·85	0·85	0·89	0·93	1·0	1·4	5°N	20·0
λ6300	1711–1717	Polarization measurements made as arc passed through zenith																		
	1719	4·1	3·1	1·9	1·8	1·7	1·7	1·7	1·7	5·0	11·7	64·4	24·0	5·2	3·5	2·8	2·0	4·1	20°S	
	172030	—	—	—	1·8	1·7	1·7	2·5	4·2	7·5	13·3	49·2	35·9	2·6	1·8	1·9	2·0	4·1	25°S	66·7
	1722	5·4	3·1	2·8	2·7	2·6	3·4	3·4	4·2	4·2	5·8	17·8	33·4	8·6	4·4	2·8	3·1	4·1	30°S	
λ5577	1723	1·11	0·74	0·65	0·20	0·20	0·19	0·19	0·19	0·37	0·37	0·75	1·32	1·56	4·65	6·70	9·42	2·60	70°S	
	1728	2·08	1·29	0·86	0·73	0·70	0·60	0·52	0·52	0·52	0·60	0·75	1·59	2·03	4·44	*	*	10·1		
	1729	—	—	—	—	0·98	0·94	0·75	0·74	0·74	0·74	0·94	1·70	2·54	4·24	9·50	17·8	12·2	75°S	21·3
λ6300	173030	—	—	—	—	—	—	—	0·83	0·83	1·7	7·6	21·0	6·9	3·5	3·7	4·1	5·4	25°S	23·9
	1734	5·4	3·1	1·9	1·8	1·7	1·7	1·7	1·7	1·7	3·3	30·5	33·4	19·0	7·1	5·5	6·0	8·1	25°S	47·0
λ5577	1735	2·97	1·49	0·86	0·73	0·78	0·76	0·75	0·93	1·12	1·49	2·62	4·54	8·0	9·29	12·7	20·8	*		
	173630	2·97	1·49	0·86	0·73	0·78	0·76	0·75	0·74	0·74	0·93	1·68	3·40	7·02	8·69	13·2	*	22·6		
λ6300	173830	2·7	2·0	1·9	1·8	1·7	1·7	1·7	1·7	2·5	7·5	26·4	24·0	6·1	4·4	5·6	5·1	5·4	25°S	40·2
λ5577	1740	2·97	1·74	1·30	1·01	0·78	0·76	0·75	0·74	0·74	0·74	0·84	1·32	1·95	3·23	12·1	*	30·4		
	1742	3·71	2·48	2·16	2·02	0·98	0·94	0·94	0·93	1·86	1·86	1·87	2·84	2·92	5·05	11·9	26·0	33·4		
λ6300	175130	1·1	0·41	0·37	0·35	0·35	0·35	0·68	0·66	0·66	1·3	2·7	8·9	15·9	4·2	3·0	3·7	2·2	40°S	
	1753	0·54	0·41	0·37	0·70	0·70	0·68	1·0	1·0	1·40	2·0	2·7	12·3	15·9	6·4	4·5	5·7	6·5	35°S	18·1
λ5577	175430	—	—	—	—	—	—	—	—	—	—	—	—	—	0·4	0·86	2·98	5·94	77°S	6·7
	1756	—	—	—	—	—	—	—	—	0·30	0·30	0·30	0·53	1·09	1·70	3·72	*	*	80°S	
λ5577	1802–1815	Zenith measurements made																		
	1818	5·19	2·98	1·94	1·62	1·56	1·51	1·31	1·30	1·30	1·30	1·50	1·32	2·14	2·83	4·10	8·93	28·9	80°S	
	181930	—	2·98	2·16	1·62	1·56	1·51	1·50	1·49	1·49	1·49	1·68	1·70	2·14	2·83	3·24	5·46	32·3	80°S	
λ6300	1821	3·8	4·5	3·4	2·5	1·7	1·0	1·0	1·0	1·0	1·7	2·7	3·1	5·5	3·7	4·1	2·8	2·7	55°S	12·7
	1822	1·6	3·7	3·4	2·8	1·7	1·7	1·4	1·3	1·7	1·7	2·4	2·4	4·1	10·6	10·8	4·5	2·7	60°S	

*Off scale

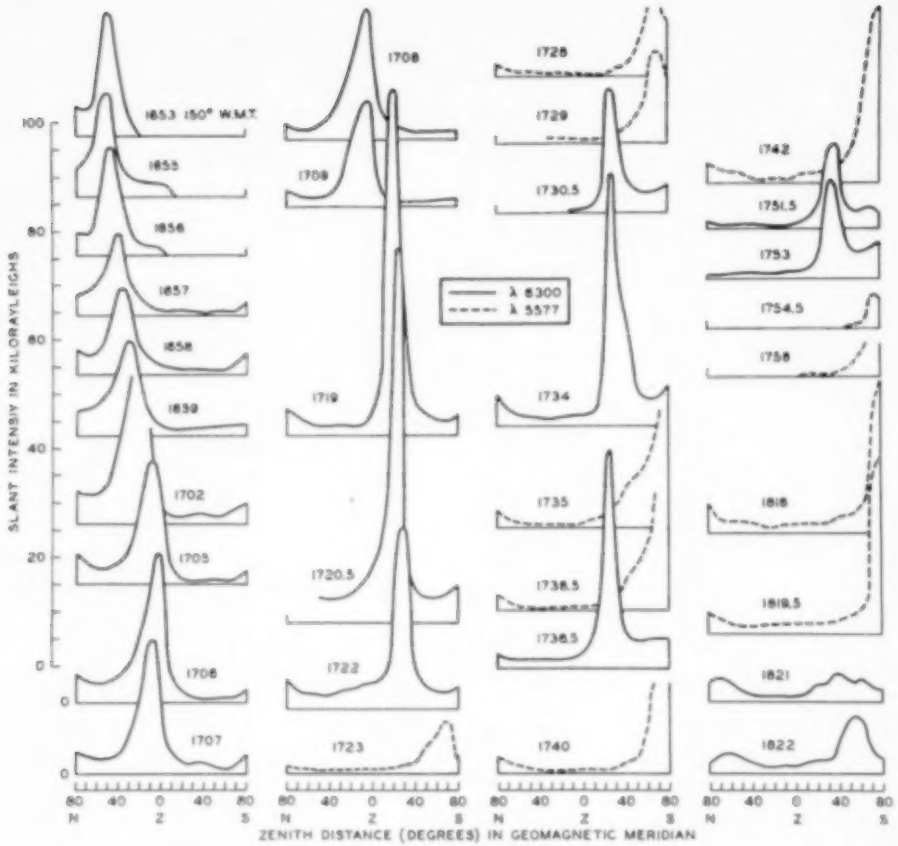


Fig. 1. The morphology of the red arc as obtained from photometric meridian scans.

E–W with the geomagnetic field and the all-sky cameras showed that the red arc was not an exception. Point Barrow is displaced 4°W in geomagnetic longitude from College. Fig. 3 shows the height determination made by plotting from each station the zenith distance in the geomagnetic meridian of the point of greatest luminosity. This angle was determined from the Barrow all-sky camera and the College λ6300 photometric data. The width of the red arc at half intensity (λ6300) as measured by the photometer (4° field of view) was approximately 15 to 20° of meridian distance. This half intensity width is represented by the error circles as drawn in the figure. The average height of the estimated greatest luminosity points is 360 km. The average deviation of these points is 20 km, and the extremes are +40 km and −30 km. The extreme deviations of the error circles are +120 km and −110 km. The heights we obtain are in keeping with the visual estimate of Elvey and Leinbach[12] who obtained 350 km for the height of a red arc at College which developed into a complete red aurora. The 1956 red arc was not sunlit, judging from its position and time of occurrence, while in the present case the arc was continually sunlit until it was far south of College. The height of the geometric shadow line in the geomagnetic meridian passing through

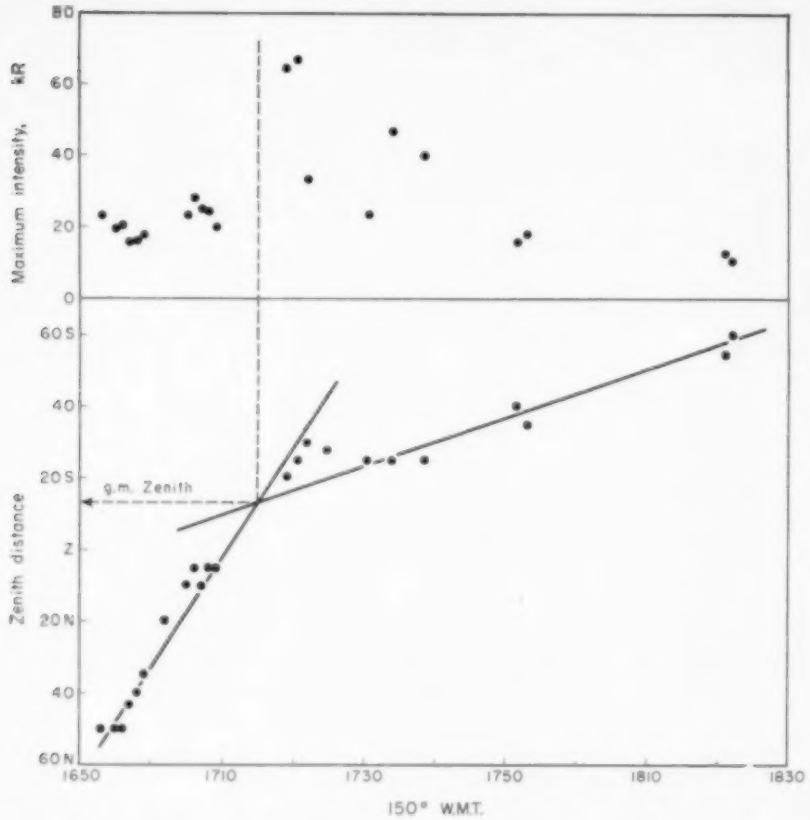


Fig. 2. The peak intensities and position of the red arc over College, Alaska.

College is indicated in Fig. 3 for various times during evening twilight.

The advent of a positive bay in the **H** component of the geomagnetic disturbance field in northern latitudes is very rare. The **H** field magnetograms from Barrow, College and Sitka, Alaska, are reproduced in Fig. 4 to show the occurrence of a positive bay concurrent with the appearance of the red arc. It is interesting to note that this magnetic disturbance did not occur simultaneously at all three locations and that the red arc was approximately in the zenith of each station during the time when **H** was returning from a positive bay.

A number of spectra were obtained while the arc moved from the northern to the southern sky at College. These are shown in Fig. 5, together with an all-sky camera frame exposed some time during each 15-min spectroscopic exposure. The spectral emission features which can be definitely associated with the red arc are the OI λλ6300-6364 lines and the N_2^+ 1 NG bands. The OI green line, λ5577, is present over the entire meridian but is not enhanced in the region of the red arc. The difference in the spectral character of the red arc compared to the low altitude arc which developed in the south is strikingly illustrated in the fourth spectrum which is reproduced in enlarged form in Fig. 6. The resolution of the instrument is insufficient for detailed analysis of the N_2^+ bands. However, it may be noted qualitatively

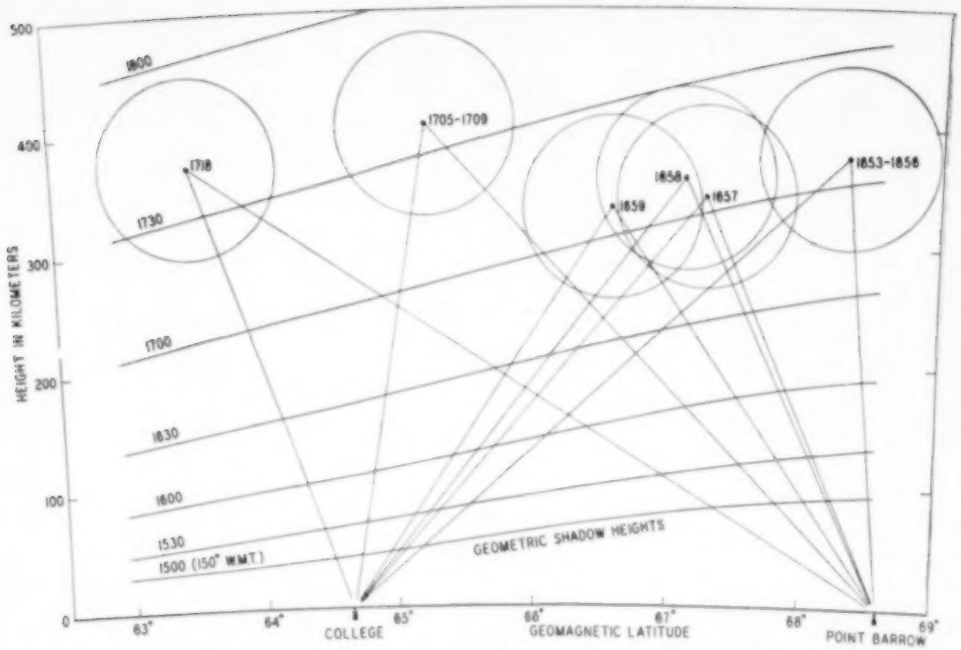


Fig. 3. The height of the red arc over College, Alaska. The height of the solid earth's shadow at various times is also shown.

that the rotational structure of these bands is considerably more developed in the red arc compared to the green arc.

Examination of the meridian spectra reveals an interesting morphology of the Balmer $\mathbf{H}_\alpha$ emission line relative to the red arc. The southward motion of the hydrogen emission, a common aurorally associated phenomenon[13], seems more rapid than that of the red arc. However, the hydrogen emission arises primarily around the 100 km level[14]. In Fig. 7 we have plotted the zenith distance of a spectral feature originating at 350 km (λ6300) against the zenith distance of another spectral feature originating at 100 km ($\mathbf{H}_\alpha$) assuming that the two emissions lie along the same field line. The solid curve indicates the expected relationship while the few available experimental points are indicated by the rectangular boxes. While good spatial resolution is available for λ6300 from the photometric data, the 15-min spectra from which the $\mathbf{H}_\alpha$ positions were deduced incorporate a good deal of smearing, in particular since the emission feature moved so rapidly.

The sodium *D* lines are enhanced in twilight and also appear to be significantly enhanced in the low altitude green arc in the south. The N_2 1 PG bands are effectively absent in the red arc and are very pronounced in the green arc. λ5577 is enhanced in the green arc. There appear several lines of mercury originating from scattered light from street lamps 2 miles distant.

THE AURORAL DISPLAY

The aurora commenced at twilight with the red arc described in the previous section. A very intense, predominantly green arc formed south of 45° zenith distance from a band which moved from east to west to span the sky by about 1720 AST. This arc then moved south beyond the horizon by 1900 AST. The red arc weakened and moved south.

At 2015 AST the intense green aurora below the southern horizon began to move northward

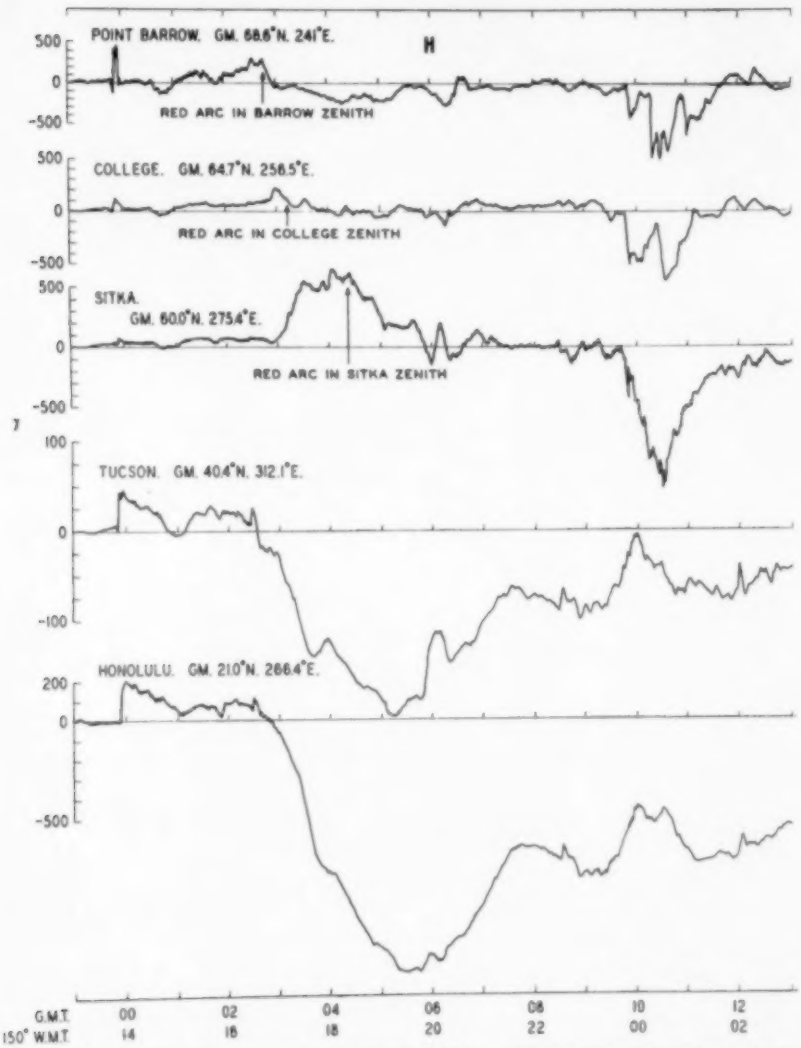


Fig. 4. The disturbance geomagnetic field at several stations for the periods 2300–0000, 27 Nov. 1959 and 0000–1300, 28 Nov. 1959. The time at which the red arc was approximately in the zenith of three stations is indicated.

and by 2020 AST reached the zenith and covered the entire sky at 2027 AST. After a short period of great activity the aurora almost disappeared until 2220 AST when a faint arc, associated with strong hydrogen emission, passed through the College zenith and moved beyond the southern horizon. Around 2345 AST intense rayed bands formed in the southern part of the sky, and covered the entire sky for about one hour beginning at 2352 AST. The post-breakup aurora which ensued for about one and one-half hours consisted of broken rayed arcs scattered throughout the sky. Thereafter the activity almost disappeared and heavy clouds moved in at 0400 AST. The sky cleared by 0515 AST, however, and faint

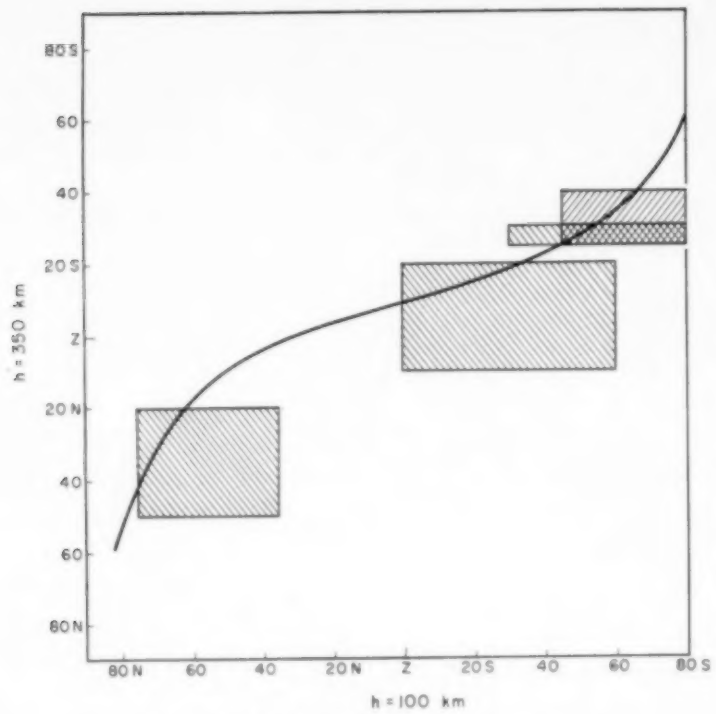


Fig. 7. Geometric correlation between emission features originating at different height levels along a field line having a dip angle of 13°. The red arc originates at 350 km while the hydrogen emission arises at the 100 km level. The zenith distances of the observations are indicated by the rectangles.

broken arcs moved up from the south to cover the sky with pulsating patches beginning about 0520 AST and lasting into morning twilight.

The great activity at College from 1950 to 2045 AST was concurrent with observations of bright aurora around midnight local time in Texas, New Mexico, South Dakota, and Minnesota[15]. However, the College aurora was far more intense after local midnight. The subvisual red arc monitored by King and Roach[16] throughout the night of 27/28 November reached the zenith at Fritz Peak (gm. lat. 49°) at about this time (0400 MST or 0100 AST). Apparently, the red arc reached its southernmost point during the time of greatest activity in the auroral zone.

Intensities in the geomagnetic zenith derived from spectroscopic data taken throughout the night are given in Table 2. The average intensity is indicated for each 15-min exposure which could be read. No readings were taken when the emission feature faded into the background or when the density was too high to be read from the calibration curve. The density of the λ5577 line was generally too great. The (0, 2) and (1, 3) bands of N_2^+ 1 NG situated at λ4709 and λ4652 respectively show a reversal of intensities from the normal behaviour during parts of the night, the reversal being most pronounced in the red arc. As would be expected, the ratio of intensities of λ6300 to λ6364, the oxygen red line doublet, was essentially constant, the experimental value being $3{\cdot}0 \pm 0{\cdot}2$, where 0·2 is the mean standard deviation of 23 readings.

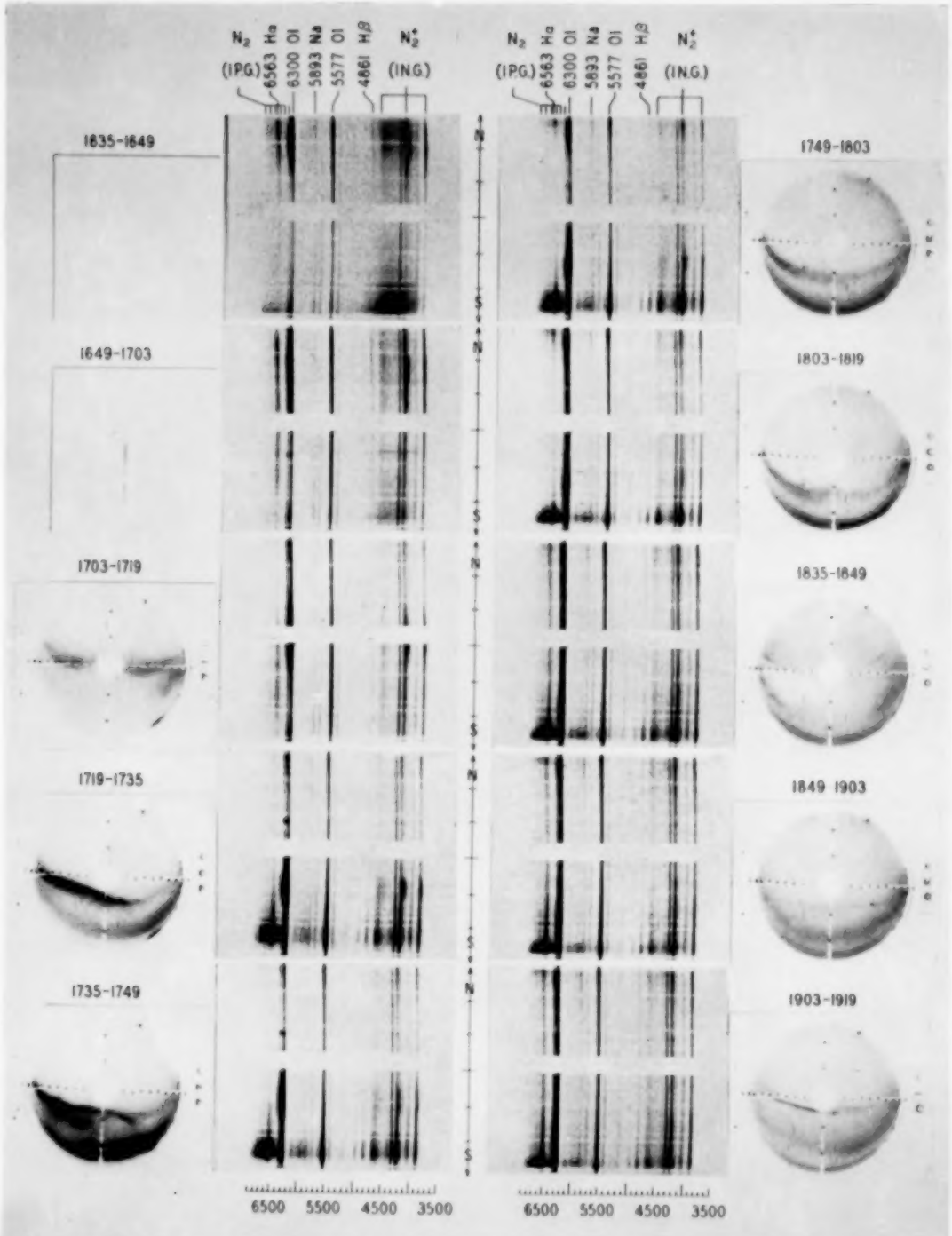


Fig. 5. College, Alaska, meridian spectrograph exposures and all-sky camera photographs for the period 1635–1919 (150° WMT) 27 Nov. 1959.

(To face page 56)

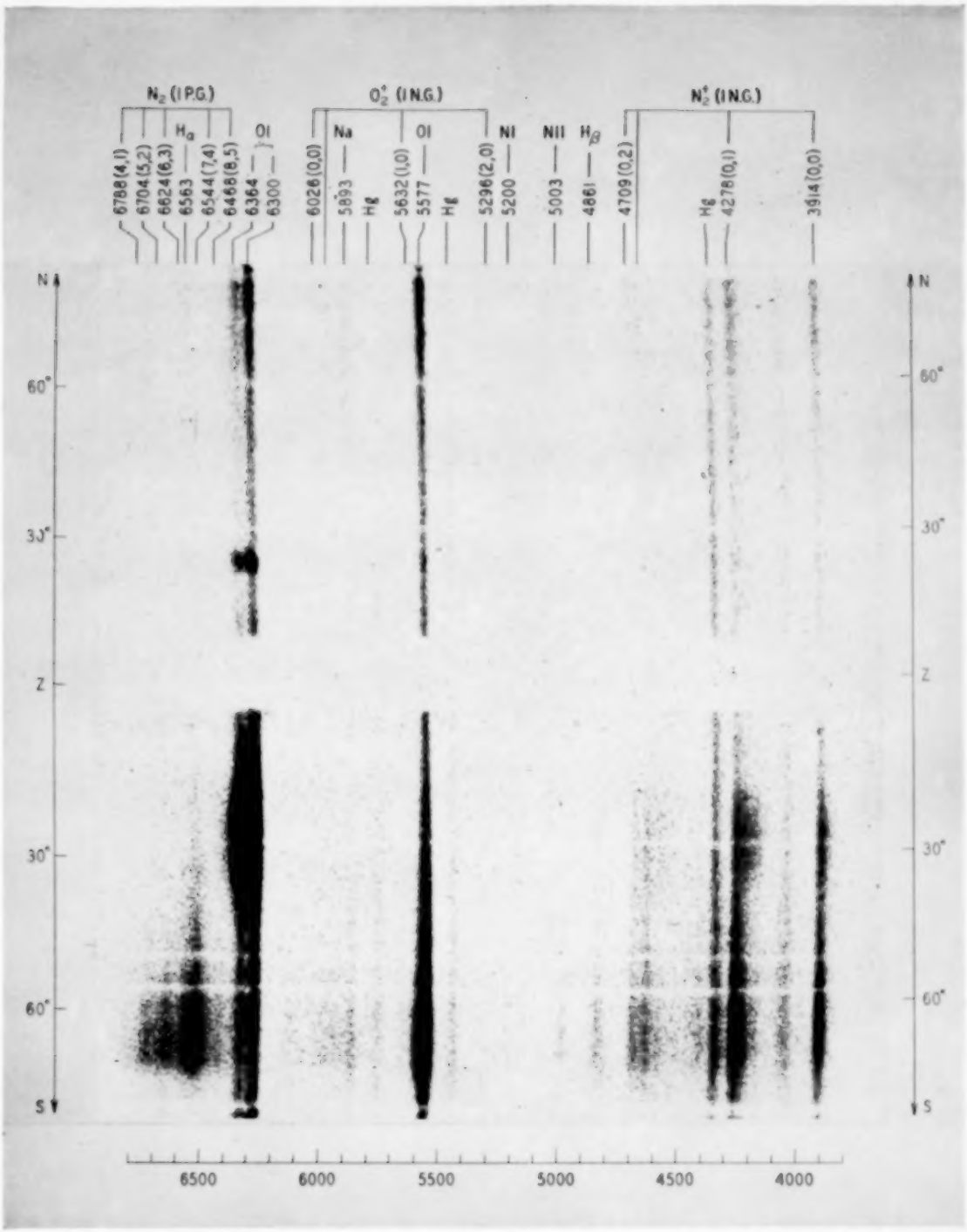


Fig. 6. Spectrum for the period 1719–1735 (150° WMT) resolving the geomagnetic meridian along various auroral emissions, and indicating the difference in spectral composition between the red arc and normal aurora.

Table 2. Zenith Intensity in Kilorayleighs—College, Alaska, 27-28 November, 1959

Time (150 WMT)	λ5577	λ4709	λ4652	λ6300	λ6364	λ6563	λ4861
1630–1645	1·29	- - - *	- - -	1·00	0·33	- - -	- - -
1645–1700	> *	0·3	- - -	1·82	0·56	0·8	0·32
1700–1715	>	0·47	0·63	>	4·3	0·83	0·51
1715–1730	2·63	- - -	- - -	4·6	1·92	0·36	- - -
1730–1745	3·8	- - -	- - -	>	2·00	0·43	- - -
1745–1800	1·92	- - -	- - -	2·64	0·96	- - -	- - -
1800–1815	3·25	- - -	- - -	4·08	1·44	0·42	- - -
1815–1830	- - -	- - -	- - -	- - -	- - -	- - -	- - -
1830–1845	5·24	0·37	0·42	4·08	1·5	0·84	0·46
1845–1900	4·33	0·23	0·23	3·54	1·2	0·87	0·3
1900–1915	4·75	0·28	0·37	4·38	1·62	0·83	0·42
1915–1930	5·03	0·32	0·33	3·23	1·12	0·8	0·66
1930–1945	4·9	0·41	0·20	2·58	0·83	0·74	0·31
1945–2000	>	0·57	0·53	3·54	1·14	0·87	0·48
2000–2015	>	0·68	0·64	>	2·22	0·72	0·66
2015–2030	>	1·28	1·29	>	>	1·82	0·98
2030–2045	>	0·48	1·48	4·54	1·55	0·63	0·48
2045–2100	>	0·44	0·44	3·45	1·16	- - -	- - -
2100–2115	>	0·36	0·40	3·24	1·08	- - -	- - -
2115–2130	>	0·41	0·40	3·42	1·1	- - -	- - -
2130–2145	>	0·22	- - -	2·72	0·9	- - -	- - -
2145–2200	5·3	- - -	- - -	2·25	0·78	- - -	- - -
2200–2215	>	0·34	0·34	2·56	0·89	- - -	- - -
2215–2230	>	0·47	0·44	2·95	0·93	0·78	0·56
2230–2245	>	0·33	0·36	3·82	1·37	1·02	0·46
2245–2300	>	0·29	0·34	>	1·7	0·82	0·5
2300–2315	>	0·47	0·46	>	1·54	0·77	0·54
2315–2330	>	0·50	0·5	>	1·63	- - -	- - -
2330–2345	>	0·55	0·5	4·18	1·36	0·53	- - -
2345–2400	>	2·32	1·89	>	2·5	0·97	0·9
2400–2415	>	1·64	1·34	>	1·52	0·54	0·51
2415–2430	>	2·72	1·98	>	1·98	1·28	1·04
2430–2445	>	2·14	1·78	>	2·44	1·83	1·24
2445–0100	>	2·26	1·69	>	2·07	2·60	1·47
0100–0115	>	0·76	0·64	4·68	1·57	1·02	0·98
0115–0130	>	0·98	0·68	3·89	1·14	1·17	0·95
0130–0145	>	0·59	0·54	3·81	1·12	0·72	0·49
0145–0200	>	0·49	0·44	4·11	1·35	0·54	0·36
0200–0215	>	0·41	0·41	3·68	1·08	0·64	0·58
0215–0230	4·75	0·26	0·31	2·48	0·71	0·58	0·42
0230–0245	5·03	- - -	- - -	2·16	0·61	0·76	0·36
0245–0300	3·44	- - -	- - -	1·86	0·56	0·6	0·36
0300–0315	2·78	- - -	- - -	1·58	0·58	0·7	0·36
0315–0330	3·44	- - -	- - -	1·56	0·62	0·74	0·31
0330–0345	2·24	- - -	- - -	1·25	0·54	0·62	0·21
0345–0400	2·33	- - -	- - -	1·22	0·42	0·43	0·31
0400–0415	2·95	- - -	- - -	1·24	0·26	0·5	- - -
0415–0430	>	- - -	- - -	- - -	>	- - -	- - -
0430–0445	>	- - -	- - -	- - -	>	- - -	- - -
0445–0500	>	- - -	- - -	- - -	>	- - -	- - -
0500–0515	>	- - -	- - -	- - -	>	- - -	- - -
0515–0530	>	0·61	0·55	2·54	0·75	- - -	- - -
0530–0545	>	0·69	0·52	2·18	0·77	- - -	- - -
0545–0600	>	0·53	0·41	1·8	0·62	- - -	- - -
0600–0615	>	0·64	0·45	1·9	0·54	- - -	- - -
0615–0630	4·64	0·48	0·42	1·84	0·56	- -	- - -
0630–0645	4·12	0·50	0·47	2·04	0·64	- - -	- - -

* > density too great to be read

- - - too faint or obscured by other emissions or too cloudy

The hydrogen emissions H_α and H_β are sufficiently far apart in the spectrum to present a serious problem in comparing their intensities. The characteristic curves of the spectroscopic film are quite different at these two wavelengths requiring a very careful calibration. From 32 spectra the average photon emission ratio H_α/H_β was found to be 1·65 with a mean standard deviation of +0·58 and −0·34. Very few values are available for this ratio. Vegard[17] quotes a (intensity) ratio of 7 for a red aurora noting that this is somewhat higher than normal. Galperin[18] and Shuiskaja[19] give values ranging from 2·8 to 3·2, while a theoretical computation by Chamberlain[20] yielded 3·34.

In general, the maxima of the emission intensities observed occurred between 1950 and 2045 AST and 2350 and 0130 AST. These maxima were simultaneous with the large polar magnetic storms of 27 November 1959 (Fig. 4), and with reports of aurora from southern U.S.A.

Acknowledgements—The observations reported herein were made while the authors were on the staff of the Geophysical Institute of the University of Alaska. Data and valuable discussions were contributed by S. Akasofu, A. Belon, T. N. Davis, and G. Romick to whom we extend our gratitude. The work was supported by a National Science Foundation grant and U.S. Air Force Contract No. AF 61(052)-131.

REFERENCES

1. C. Störmer, *The Polar Aurora*, Oxford University Press, Oxford (1955).
2. C. T. Elvey, *Proc. Nat. Acad. Sci., Wash.* **43**, 63 (1957).
3. K. C. Clark and G. J. Romick, *An Optics Field Site for Auroral Studies*. Geophysical Institute, Univ. of Alaska, Scient. Rep. No. 1, NSF Grant Y/114/177 (1958).
4. R. B. Dunn and E. Manring, *J. Opt. Soc. Amer.* **46**, 572 (1956).
5. L. R. Megill, *J. Atmos. Terr. Phys.* **17**, 276 (1960).
6. F. E. Roach, *National Bureau of Standards* Annual Report No. 5591 (1958).
7. C. S. Deehr, *A Spectrophotometric Study of the Aurora of 27 November 1959 at College, Alaska*, M.S. Thesis, Univ. of Alaska (1961).
8. K. C. Clark and G. J. Romick, *J. Opt. Soc. Amer.* **49**, 141 (1959).
9. C. S. Deehr, letter submitted to *J. Opt. Soc. Amer.* (1961).
10. C. T. Elvey and A. E. Belon, *Description of the all-sky camera* Geoph. Inst. Univ. of Alaska, Scient. Rep. No. 1, NSF Grant Y/1.1/44 (1957).
11. R. A. Duncan, *Planet. Space Sci.* **1**, 112 (1959).
12. C. T. Elvey and H. Leinbach, *Sky and Telesc.*, **15**, 329 (1956).
13. M. H. Rees, A. E. Belon and G. J. Romick, *Planet. Space Sci.* **5**, 87 (1961).
14. M. H. Rees, *Mém. Soc. Roy. Sci. Liège* **20**, 609 (1961).
15. C. Gartlein, *U.S. Visual Observations, News Letter No. 18*, Ithaca, New York (1959).
16. G. A. M. King and F. E. Roach, *J. Res. Nat. Bur. Stand.* **65D**, 129 (1961).
17. L. Vegard, *Geofys. Publ.* **19**, No. 4 (1955).
18. G. I. Galperin, *Planet. Space Sci.*, **1**, 57 (1959).
19. F. K. Shuiskaja, *Astr. J. Moscow*, **37**, 186 (1960).
20. J. W. Chamberlain, *Astrophys. J.* **120**, 360 (1954).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 59–67. Printed in Great Britain.

EXCITATION OF HIGH ALTITUDE RED AURORAL ARCS

M. H. REES
Department of Applied Mathematics, Queen's University, Belfast*

(*Received* 27 *June* 1961)

Abstract—The excitation of high altitude red auroral arcs is discussed using extensive observational data obtained for one such arc which occurred at College, Alaska, in the evening of 27 November 1959. The arc cannot be explained solely by thermal effects or local accelerating fields. Charged particles incident on the atmosphere must be invoked. It is argued that these particles are probably protons rather than electrons. The mean proton energy is estimated to be a few keV. A proton flux of 1×10^{10} cm^{-2} sec^{-1} is required to account for the maximum observed luminosity.

INTRODUCTION

Photographic, photometric, and spectroscopic observations on a high altitude red arc made in Alaska during the evening of 27 November 1959 have recently been described in some detail by Rees and Deehr[1]. Based on these observational data the possibilities of excitation by locally accelerated ionospheric electrons, and by electrons or protons incident on the atmosphere are examined. In spite of uncertainties arising from the crudeness of some of the data and lack of precise reaction rates arguments can be advanced favouring incoming low-energy protons as being principally responsible.

The atmospheric model (Table 1) given by Nicolet[2] is used in all computations.

Table 1. Model Atmosphere[2]

km	n(O) cm^{-3}	n(N_2) cm^{-3}	n(O_2) cm^{-3}
200	$1{\cdot}34 \times 10^9$	$6{\cdot}18 \times 10^9$	$1{\cdot}06 \times 10^9$
250	$7{\cdot}05 \times 10^8$	$1{\cdot}96$	$3{\cdot}17 \times 10^8$
300	$4{\cdot}08$	$6{\cdot}56 \times 10^8$	$1{\cdot}06$
350	$2{\cdot}38$	$2{\cdot}23$	$3{\cdot}60 \times 10^7$
400	$1{\cdot}40$	$7{\cdot}71 \times 10^7$	$1{\cdot}24$
450	$8{\cdot}27 \times 10^7$	$2{\cdot}70$	$4{\cdot}36 \times 10^6$
550	$2{\cdot}97$	$3{\cdot}48 \times 10^6$	$5{\cdot}61 \times 10^5$
650	$1{\cdot}10$	$4{\cdot}75 \times 10^5$	$7{\cdot}67 \times 10^4$
750	$4{\cdot}16 \times 10^6$	$6{\cdot}85 \times 10^4$	$1{\cdot}11$

KEY OBSERVATIONAL DATA

With details given by Rees and Deehr[1] only the most pertinent quantitative observations are listed with some qualifying comments occasionally added.

The principal spectral features were the λλ6300-6364 lines of OI, which gave the arc its red colour, and the N_2^+ 1N bands. The hydrogen lines could have been associated with the red arc if allowance for a different emission height were made[1]. However, the observations were not sufficiently refined to prove this association. The maximum photon intensity of OI λ6300, $4\pi\mathcal{I}(6300)$, was 66·7 kR (kilorayleighs†. It is thought to have lasted less than ten minutes. The highest value of $4\pi\mathcal{I}(6300)/4\pi\mathcal{I}(5577)$ recorded photometrically was 25. The average intensity of the λ4709 band of N_2^+ deduced from a 15 min spectrogram was 0·47 kR. During this interval the red arc moved from 20° north to 20° south zenith distance and showed changes in the (photometric) intensity of λ6300 by as much as a factor of three. The average (spectroscopic) intensity of the hydrogen line H_β was 0·51 kR during this period.

Throughout its life over College, Alaska, the arc was situated in the sunlit atmosphere at an average altitude of 350 km; its width was of the order of 100 km; it was aligned approximately along parallels of the geomagnetic field.

* *Present address*: Department of Physics, University of Colorado, Boulder, Colorado.

† The photometer selectively monitored the λ6300 line of the doublet.

Serious error should not be introduced by assuming that the emission occurred uniformly from an altitude of (350-H) km to an altitude of (350+H) km, H being the scale height*.

The red arc necessitates a mechanism which yields a high photon intensity of the OI λλ6300-6364 lines, a moderate photon intensity of the N_2^+ 1N bands but virtually none of the OI λ5577 line and the N_2 1P bands. Table 2 shows the relevant threshold energies and the required rates of formation of the excited species†. The pattern revealed is peculiar to the red arc.

Table 2.

Spectral feature	Threshold energy eV	Required rate of formation $cm^{-3}\ sec^{-1}$
O (^1D)	1·97	$3{\cdot}6\times10^3$
O (^1S)	4·17	$<1{\cdot}5\times10^2$
N_2 $B^3\Pi$	7·5	$<4\ \times10^1$
N_2^+ $X^2\Sigma$	15·6	$1\ \times10^3$

EXCITATION BY IONOSPHERIC ELECTRONS

Energetic ionospheric electrons may be produced by thermal agitation or by acceleration in an electric field associated with a current flowing in the high atmosphere. Both processes yield a selective energy distribution of electrons which could account for the large ratio of λ6300 to λ5577 and the absence of the N_2 1P bands. Neither process can account for the ionization of N_2. Seaton[3] mentioned the possibility that N_2^+ ions may be produced at lower levels (presumably by particle bombardment) and transported to greater heights. The rapid latitude shift of the red arc makes this possibility seem very unlikely in the present case.

If the formation of the N_2^+ ions is attributed to some as yet unidentified mechanism one may investigate to what extent ionospheric electrons can explain the OI radiation.

The photon intensity of the λ6300 line may be expressed by

$$4\pi\mathscr{I}\,\mathrm{I}\,(6300)=0{\cdot}75\,s_{12}\int_{350-H}^{350+H} n\,(0)\,n\,(e)\,\mathrm{d}h \qquad (1)$$

in which 0·75 is the fraction of λ6300 quanta emitted in the $^1D-{}^3P$ transition, s_{12} is the excitation coefficient of the ^1D term, and $n\,(0)$ and $n\,(e)$ are the concentrations of atomic oxygen and of electrons respectively.

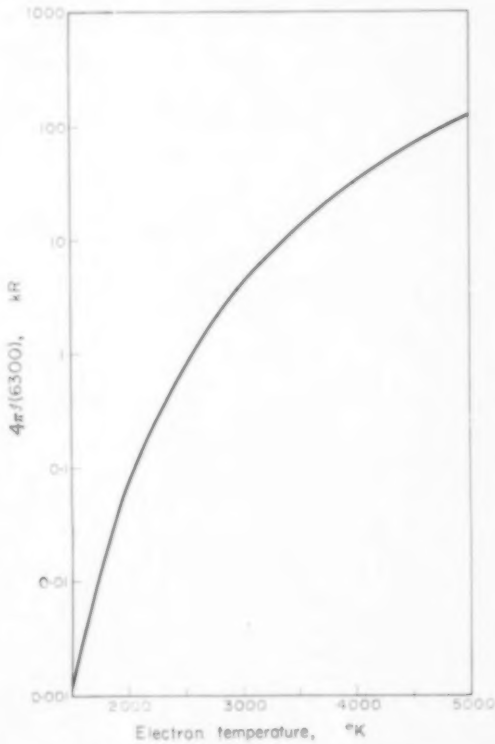


Fig. 1. Photon emission rate of λ6300 for excitation by thermal electrons from an altitude of (350—H) to (350+H) km with $n(e)=1\times10^6 cm^{-3}$.

The expected photon intensity is shown as a function of electron temperature in Fig. 1, the excitation coefficient for thermal electrons calculated by Seaton[4] for a Maxwellian distribution

* An unsuccessful attempt was made to construct the height-luminosity profile, but a unique distribution satisfying the observational data on the arc at four different zenith distances could not be found.

† It is assumed that resonance scattering of solar radiation by N_2^+ ions is the source of the 1N bands. The relevant threshold energy is thus the ionization potential of N_2. The required rate of formation of N_2^+, which is naturally what is given in the last column, was estimated on the basis that 0·47 kR of λ4709 originating over two scale heights needs an N_2^+ concentration of $1{\cdot}5\times10^3\ cm^{-3}$.

being adopted, and the electron density being taken to be 1×10^{6} cm^{-3}. Clearly, 66·7 kR would require an unacceptably high temperature. Recalling Mulyarchick's[5] measurements of a temperature (during a red aurora) in excess of 2500°K (from Doppler broadening of the λ6300 line) it may be argued that thermal electrons may account for the diffuse, formless red background glow which frequently is present during very bright aurorae. The excitation coefficient for the ^1S term, s_{13}, is extremely small and no λ5577 emission would result. It is difficult, however, to conceive a mechanism which can produce high temperature electrons in a narrow, field-aligned region even for a short period of time.

Equation (1) also expresses excitation by an electric discharge, s_{12} representing the appropriate coefficient. A detailed theory has been given by Chamberlain[6] who computed s_{12} and s_{13} for various values of electric field and mean free path. At 350 km a field of 2 V/km would make $4\pi\mathcal{J}(6300)/4\pi\mathcal{J}(5577)\geqslant 25$ and if $n(e)$ were 4×10^{5} cm^{-3} $4\pi\mathcal{J}(6300)$ would then be as much as 66·7 kR. The mechanism may be partly responsible for the red arc.

EXCITATION BY INCIDENT ELECTRONS

Normal aurorae are produced principally by incident energetic electrons, and secondary electrons are probably responsible for most of the excitation of OI[7, 8].

If the mechanism for the red arc were basically the same as that for normal aurorae, the difference in the observed spectrum should be attributable to the difference in emission heights. At 350 km collisional deactivation of the ^1D levels may be neglected giving for the ratio of the excitation rates[8],

$$\frac{S_{12}}{S_{13}}=0{\cdot}94\left[\frac{4\pi\mathcal{J}(6300+6364)}{4\pi\mathcal{J}(5577)}-1\right] \qquad (2)$$

which, for $4\pi\mathcal{J}(6300)/4\pi\mathcal{J}(5577)=25$ requires $S_{12}/S_{13}\simeq 30$. This is more than twice the value obtained by Seaton[9] and Omholt[7] from considerations on excitation and deactivation of OI in normal aurorae. If excitation of N_2 to the $A^3\Sigma$ and $B^3\Pi$ states is taken into account Omholt obtains S_{12}/S_{13} as great as required by (2) but this could not apply in the red arc since the N_2 1P bands are absent in the spectrum.

Therefore, a height effect alone does not support the hypothesis. It would be necessary to invoke an energy distribution of secondary electrons different from that obtained in the usual auroral ionization processes. There is little justification for such an assumption.

No doubt, for penetration to the 350 km level the energy of the incident primary electrons need not be large, with a consequent reduction in the number of ion pairs formed per incident electron. A rough estimate of the required particle flux gives about 2×10^{11} electrons cm^{-2} sec^{-1} if excitation of the red arc were attributable to this mechanism.

EXCITATION BY INCIDENT PROTONS

A crude extrapolation of available range—energy data[10] yields only about 200 eV for the energy of protons following a rectilinear path in Nicolet's[2] model atmosphere to penetrate to the 350 km level. Even considering the effect of spiralling of the protons, their average energy may be no more than 1 keV in which case the $H-H^+$ beam will consist of more than 95 per cent neutral hydrogen atoms. However, repeated capture and ionization processes require consideration of several possible primary reactions. For the present case the important reactions are:

$$H^+ + O \longrightarrow H + O^+ \text{ and } H + O \longrightarrow H + O^+ + e \qquad (3)$$

$$H + O(^3P) \longrightarrow H + O(^1D) \qquad (4)$$

for atomic oxygen, and

$$H^+ + N_2 \longrightarrow H + N_2^+ \qquad (5)$$

$$H^+ + O_2 \longrightarrow H + O_2^+ \qquad (6)$$

for N_2 and O_2 respectively. The N_2^+ and O_2^+ ions disappear by dissociative recombination,

$$N_2^+ + e \longrightarrow N + N \qquad (7)$$

$$O_2^+ + e \longrightarrow O + O \qquad (8)$$

It is arbitrarily assumed that reaction (8) gives rise, on the average, to one atom in the ^1D levels.

Atomic oxygen ions disappear principally by ion-atom exchange,

$$O^+ + N_2 \longrightarrow NO^+ + N \quad (9)$$

$$O^+ + O_2 \longrightarrow O_2^+ + O \quad (10)$$

Reaction (9) is followed by,

$$NO^+ + e \longrightarrow N + O \quad (11)$$

Oxygen atoms produced by (11) cannot end in the ^1S level.

Statistical equilibrium between the various reactions cannot be established due to the rapid intensity variations and the southward motion of the red arc observed at College. It is shown in the appendix that diffusion of the $O(^1D)$ atoms becomes important for thin arcs at high altitude. Their long lifetime enables the atoms to diffuse appreciably before emission occurs with the effect that an initially narrow λ6300 feature will eventually spread out. On the other hand, diffusion of the ions, even for the long-lived O^+ species, is effectively impeded by the geomagnetic field since diffusion outward from the arc will be transverse to the lines of force.

The continuity equation for $O(^1D)$ atoms (labelled O′) is,

$$\frac{\partial n(O')}{\partial t} = D\frac{\partial^2 n(O')}{\partial x^2} + k_8 n(e)\, n(O_2^+) + k_{11} n(e)\, n(NO^+) + k_4 n(O)\, n(H) - A_{21} n(O') \quad (12)$$

in which n's refer to the concentrations of the various species, D is the diffusion coefficient (diffusion occurring in the horizontal direction), the k's are the reaction rates (subscripts referring to the equations giving the process), and A_{21} is the spontaneous transition probability for radiation of the λ6300 line. Diffusion is examined in the appendix, while the reaction rates are listed and discussed in Table 3.

In view of the rapid rise and decay of the luminosity at one point in the sky (due to the rapid southward shift of the arc) the actual situation is represented to a sufficient approximation by taking the concentration of protons in a given latitude strip of semi-width l to be constant and equal to $n(H-H^+)$ for $0 \leqslant t \leqslant T$ and zero at other times.

Solving (12) for $0 \leqslant t \leqslant T$ yields that at the centre of the arc

$$n(O') = B\int_0^t e^{-A_{21}\tau}\,\mathrm{erf}\,\frac{l}{2\sqrt{D\tau}}\,d\tau - Ce^{-Et}\int_0^t e^{(E-A_{21})\tau}\,\mathrm{erf}\,\frac{l}{2\sqrt{D\tau}}\,d\tau + Fe^{-Gt}\int_0^t e^{(G-A_{21})\tau}\,\mathrm{erf}\,\frac{l}{2\sqrt{D\tau}}\,d\tau + Ie^{-Jt}\int_0^t e^{(J-A_{21})\tau}\,\mathrm{erf}\,\frac{l}{2\sqrt{D\tau}}\,d\tau \quad (13)$$

where

$$B = [k_4 n(O) + k_6 n(O_2) + k_5 n(O)]\, n(H-H^+)$$

$$C = k_5 n(O)\, n(H-H^+), \quad E = k_9 n(N_2) + k_{10} n(O_2)$$

$$F = \frac{k_9 n(N_2)\, C}{G}, \quad G = k_{11} n(e), \quad J = k_8 n(e)$$

Table 3. Reaction Rates

Reaction rate $cm^3\ sec^{-1}$	Reference	Comments
$k_3 = 2{\cdot}5\times10^{-8}$	17	Inverse process was investigated; value at same relative energy of motion is used.
$k_4 = 4\times10^{-9}$	18	No data available; maximum cross-section supposed same for electron and heavy particle excitation.
$k_5 = 1{\cdot}1\times10^{-8}$	19	Should be reliable.
$k_6 = 2\times10^{-8}$	20	Data available only to proton energy as low as 4 keV.
$k_7 = 6\times10^{-7}$	21	May be smaller; identification of ions is questionable.
$k_8 = 3\times10^{-8}$	22	These values provide a consistent picture of normal ionospheric processes.
$k_9 = 1\times10^{-13}$	22	
$k_{10} = 2\times10^{-13}$	22	
$k_{11} = 3\times10^{-9}$	22	
$A_{21} = 6{\cdot}9\times10^{-3}$		

$$I=\frac{k_{10}n(O_2)C}{J}-k_6 n(O_2)\,n(H-H^+).$$

The average time required for the luminosity to reach a maximum was approximately six min. Numerical evaluation of (13) at $t=T=360$ sec for the maximum observed brightness (66·7 kR) gives the required concentration of $n(H-H^+)$ to be $2{\cdot}3\times10^3$ cm^{-3}. The associated ion densities under these conditions are listed in Table 4.

Table 4. Ion densities at 66.7 kR for excitation by incident protons

Ion	Concentration cm^{-3}
n (H–H+)	$2{\cdot}3\times10^3$
n (O+)	$5{\cdot}0\times10^6$
n (N_2^+)	$2{\cdot}2\times10^3$
n (NO+)	$1{\cdot}2\times10^4$
n (O_2^+)	$8{\cdot}0\times10^3$

The decay of the luminosity for $t\geqslant T$ may be determined from equation (13) by subtracting from each term in the equation the same term in which all t's have been changed to $(t-T)$. The rise and decay of the luminosity with the assumption of a constant particle flux lasting for six min is shown in Fig. 2. The residual luminosity tail falling off very slowly with time is due to the very slow recombination of the O^+ ions formed by the incoming stream. The relative importance of the direct excitation process (4) compared to the indirect processes (8), (11) depends on the magnitude of the respective reaction rates. For the parameters adopted in the present case the contributions to the maximum luminosity by the two processes are about the same.

The experimental points plotted in Fig. 2 were obtained by averaging the decay observed at several zenith distances. Since the intensities were different at various zenith distances the data are presented in normalized form. Exact agreement with the model would not be expected, but the experimental decay points should all fall above the computed curve since a more rapid decay would not be possible. Experimental points corresponding to the build-up of brightness would be expected to fall below the computed curve since a step function was assumed for the incident flux.

Using the N_2^+ concentration given in Table 4 leads to an expected intensity of 1·8 kR for the λ4709 band from resonance scattering, while

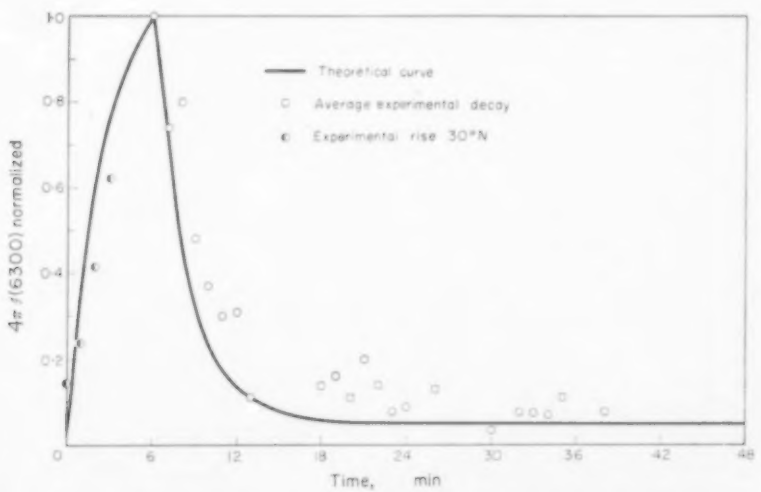


Fig. 2. Normalized photon emission rate of λ6300 for excitation by a constant proton flux of 6-min duration. The experimental points for the decay of the luminosity are averaged from several photometric scans.

direct excitation contributes an additional 0·5 kR. This emission lasts for the 6-min interval during which the proton flux impinges on a given volume. The expected emission is therefore approximately twice that deduced from the 15-min exposure.

If the average energy of the incident protons is 1 keV and each proton undergoes about ten charge exchanges the flux corresponding to the maximum observed luminosity is about 1×10^{10} cm^{-2} sec^{-1}.

With the above proton flux some hydrogen emission would be expected to originate in the volume of the red arc. Using appropriate excitation cross-sections for H_β[11] an intensity of 30 rayleighs is predicted which is too low to have been detected with the instrumentation employed. The observations show[1] that hydrogen radiation may be related to the red arc, but the emission originates much lower in the atmosphere. This may be attributed to the large range of proton energies required to penetrate to different levels of the atmosphere. General theoretical considerations[12] indicate that the fraction of captures into excited states of hydrogen is much smaller at low energies. The observed intensity of H_β, 0·51 kR, originating at 100 km would require a flux of approximately 2×10^7 protons cm^{-2} sec^{-1} of 70 keV energy, about 500 times smaller than the *particle* flux at 350 km but only seven times smaller in terms of energy flux.

RAY STRUCTURE IN THE RED ARC

A puzzling aspect of the red arc phenomenon is the visual observation by an experienced auroral observer (Mr. C. S. Deehr) of the occurrence of *red* ray structure in the arc with east-west motion*. Regardless of whether this represents actual movement of individual rays or is merely an apparent motion caused by a wave of excitation sweeping along the arc, the observation implies at least a short-lived existence of very narrow structures.

If the red colour is attributed to the λλ6300-6364 radiation diffusion of the excited atoms at 350 km is rapid and a narrow arc or ray cannot be maintained long with a continuous source of excitation. The width of the rays was not measured; features as narrow as 250 m have been reported[14] but a greater width would be more favourable to the observation. Assuming the width of the ray to be 2 km initially the luminosity rapidly decreases as shown in Fig. 3. Either the incoming particle flux is not continuous but impinges on the atmosphere with a periodicity of the order of seconds or the ray structure is a consequence of the electrodynamics of the arc itself. Recalling that the incident proton flux produced more than the required N_2^+ ions it may be speculated that part of the λ6300 radiation may be produced by the discharge mechanism which may also account for the ray structure in some manner as yet unexplained.

The possibility that the redness was due to the Balmer H_α line is not supported by the observations; the intensity was too low to expect visual response while the spatial correlation of hydrogen has already been mentioned.

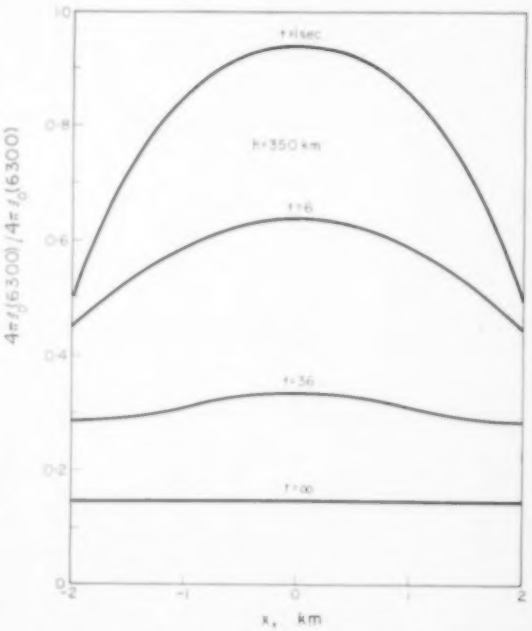


Fig. 3. The effect of diffusion in smearing out the luminosity of an initially discrete λ6300 feature at the 350 km level.

* Red ray-streamers at great altitudes have also been observed in Norway.[13]

DISCUSSION

Incident protons or electrons are needed to account for the observed intensity of N_2^+ 1N bands. Considerable heating should ensue from the incident flux, and probably more than half of the energy brought in is eventually converted into heat locally producing an increase in the temperature*. Heat will be lost by conduction and by convection. The effect of conduction can be estimated. If the maximum flux lasted as long as six min the temperature in the centre of the arc only rises from an initial value of 1580°K to about 2000°K. The effect of convection was not considered in this estimate though it may actually be the dominant heat loss mechanism. Independent observations support some increase in temperature. Vallance Jones and Hunten[15] derived a kinetic temperature in excess of 2000°K in the 400 to 500 km region from measurements of the rotational and vibrational intensity distributions of N_2^+ 1N bands in sunlit auroral rays. The spectra of the red arc[1] qualitatively show the same high rotational development of these bands indicative of a high temperature, though the resolution is far too low for quantitative analysis.

To summarize, it appears that thermal ionospheric electrons or incident energetic electrons cannot account for the red arc. No objections are found to the hypothesis that an incoming low energy proton stream is responsible. A flux of the order of 10^{10} cm^{-2} sec^{-1} for low energy protons is not inconsistent with values obtained by Obayashi[16] from analysis of magnetic storm pulsations. It is concluded that protons play the primary role in producing the red arc. However, the ray structure associated at times with the arc suggests that an electric field may be a secondary consequence which, at the same time, could produce some of the λ6300 emission.

Acknowledgements—It is a pleasure to thank Professor D. R. Bates for numerous informative and stimulating discussions throughout this investigation. I am indebted to Professor A. Dalgarno for suggestions regarding diffusion coefficients and other matters, and to Dr. R. C. McCarrol for help in computer programming. The work was supported through U.S. Air Force Grant AF-EO ARDC 61-16.

* Energy not absorbed *in situ* includes that fraction which is ultimately converted to radiation not absorbed locally, such as the N_2^+ bands and the OI lines, and some of the magnetic dipole radiation of the ground term of OI in the far infrared region of the spectrum.

APPENDIX

The excited oxygen atoms are produced in an arc, a field aligned sheet of thickness $2l$ extending along a parallel of geomagnetic latitude. The atoms will, of course, diffuse in all directions, however, diffusion along the arc and vertically within it will not result in a broadening effect of the luminosity. A one-dimensional diffusion problem results.

Let A_0 be the constant rate of production of excited atoms per unit volume and per unit time and let the excited atoms disappear by radiative de-excitation only, with a transition probability A_{21}. The number of excited atoms per unit volume at any point within the arc and at any time t is

$$\frac{\partial n(\mathrm{O}')}{\partial t} = D\frac{\partial^2 n(\mathrm{O}')}{\partial x^2} + A_0 - A_{21} n(\mathrm{O}')$$

where the notation has already been defined. The initial conditions are

$$n(\mathrm{O}') = 0 \text{ at } t = 0$$

and the boundary conditions are

$$\left.\frac{\partial n(\mathrm{O}')}{\partial x}\right]_{x=0} = 0, \quad t > 0, \quad -l < x < l$$

The solution is

$$n(\mathrm{O}') = \frac{A_0}{2}\left[\int_0^t e^{-A_{21}\tau} \operatorname{erf}\frac{l-x}{2\sqrt{D\tau}}\,d\tau + \int_0^t e^{-A_{21}\tau} \operatorname{erf}\frac{l+x}{2\sqrt{D\tau}}\,d\tau\right]$$

where the integrals must be evaluated numerically.

The physical quantity of interest is the effect of diffusion in decreasing the emission rate of λ6300 in the arc (by spreading the luminosity over a wider area). The steady state ratio of the emission rates from the centre of the arc

with and without diffusion is given by

$$\frac{4\pi\mathcal{I}(6300)}{4\pi\mathcal{I}(6300)}=A_{21}\int_0^\infty e^{-A_{21}t}\,\mathrm{erf}\frac{l}{2\sqrt{Dt}}dt=1-e^{-l\sqrt{\frac{A_{21}}{D}}}$$

Neither experimental nor theoretical data are available for the specific case of diffusion of $O(^1D)$ atoms in atomic oxygen and molecular nitrogen. For various combinations of gases Dn ranges between 2×10^{18} and 2×10^{19} cm^{-1} sec^{-1} with a temperature dependence varying between $T^{\frac{1}{2}}$ and T^{1}. Following the recommendation of Professor A. Dalgarno the value of $Dn=5\times10^{18}$ cm^{-1} sec^{-1} was adopted and D was computed using the model atmosphere of Table 1. The effect of diffusion in reducing the steady state emission rate in the centre of an arc is shown in Fig. 4 for various values of the parameters and a constant production rate of excited atoms.

REFERENCES

1. M. H. Rees and C. S. Deehr, *Plan. Space Sci.* This issue, 49.
2. M. Nicolet, *Space Research* (Ed. H. Kallman Bijl) North Holland Publ. Co., Amsterdam (1960).
3. M. J. Seaton, *The Airglow and the Aurorae* (Ed. E. B. Armstrong and A. Dalgarno) p. 225. The Pergamon Press, London (1955).
4. M. J. Seaton, *ibid*, p. 289.
5. T. M. Mulyarchick, *Spectral, Electrophotometrical and Radar Researches of Aurorae and Airglow*, No. 1, 30 (1959).
6. J. W. Chamberlain, *The Airglow and the Aurorae* (Ed. E. B. Armstrong and A. Dalgarno) p. 206, The Pergamon Press, London (1955).
7. A. Omholt, *Geofys. Publ.* **21**, No. 1 (1959).
8. D. R. Bates, *Physics of the Upper Atmosphere* (Ed. J. A. Ratcliffe) p. 297, Academic Press, New York and London (1960).
9. M. J. Seaton, *J. Atmos. Terres. Phys.* **4**, 295 (1954).
10. H. H. Landolt-Börnstein, *Zahlenwerte und Funktionen*, Springer Verlag, Berlin (1952).
11. N. P. Carleton and T. R. Lawrence, *Phys. Rev.* **109**, 1159 (1958).
12. D. R. Bates and A. Dalgarno, *Proc. Phys. Soc.*, **A66**, 972 (1953).
13. L. Vegard, *Geofys. Publ.*, **19**, No. 4 (1955).
14. C. T. Elvey, *Proc. Nat. Acad. Sci.*, **43**, 63 (1957).

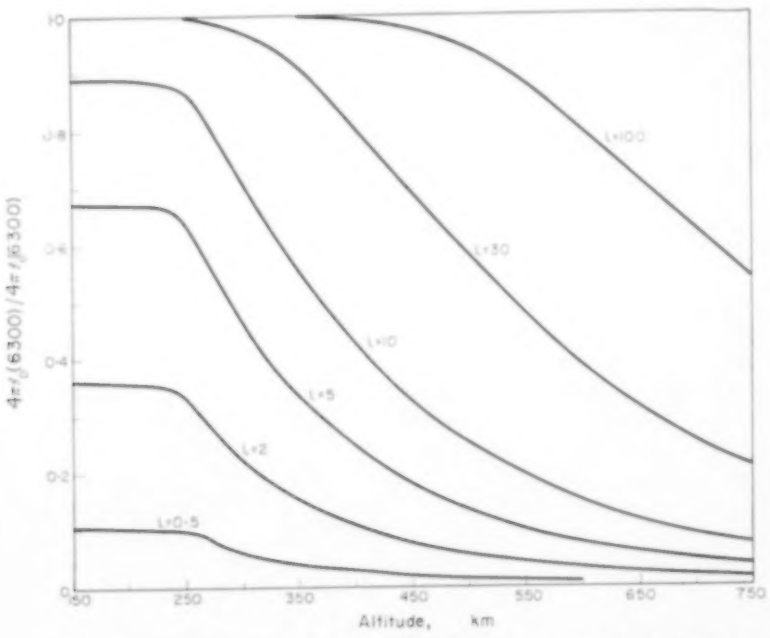


Fig. 4. The effect of diffusion in reducing the luminosity of λ6300 at the centre of an arc of width 2 l for steady state conditions.

15. A. Vallance Jones and D. M. Hunten *Canad. J. Phys.* **38,** 458 (1960)
16. T. Obayashi, *Rep. Ionosp. Res. Japan,* **12,** 301 (1958).
17. A. Dalgarno, *Ann. Géophys.,* **17,** 16 (1961).
18. D. R. Bates and M. Nicolet, *J. Atmos. Terr. Phys.* **18,** 65 (1960).
19. R. F. Stebbings, W. L. Fite and D. G. Hummer, *J. Chem. Phys.* **33,** 1226 (1960).
20. H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena* Oxford University Press (1952).
21. H. B. Gilbody and J. B. Hasted, *Proc. Roy. Soc.* **A238,** 334 (1956).
22. S. K. Allison, *Rev. Mod. Phys.,* **30,** 1137 (1958)

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 68–70. Printed in Great Britain.

THE MEASUREMENTS OF RANDOM RADIO FREQUENCY FIELDS BY THE USE OF SHORT ELECTRIC AND MAGNETIC DIPOLES

J. HUGILL
Mullard Radio Astronomy Observatory, University of Cambridge

(*Received* 15 *August* 1961)

Abstract—This paper is concerned with the relations between measured noise voltages or currents at the terminals of short antennas and the temperature of the radiation field to which the antenna is exposed. The variations in these relations due to the presence of a very simple ionosphere surrounding the antenna is considered. The theory is applied to the experimental measurements of sky brightness at 3·8 Mc/s by Molozzi, Franklin and Tyas.

INTRODUCTION

Because of the earth's ionosphere it is impossible to measure the radiation from outer space at frequencies much below 10 Mc/s at the surface of the earth. In order to extend the measurements of background radiation from the sky to lower frequencies it is necessary to make use of rockets and earth satellites which can penetrate above the reflecting layers of the ionosphere. The usual technique for measuring the brightness temperature of the sky at high frequencies is the measurement of the power developed in the radiation resistance of a matched aerial system whose radiation pattern is directed at the point of interest.

For wavelengths of about 300 metres it is impracticable to use a resonant electric dipole on a rocket or satellite; instead a short electric dipole connected to a high impedance receiver acting as a voltmeter can be used. This will in effect measure the electric field strength of the radiation field. Similarly a magnetic loop dipole will provide a current proportional to the magnetic field strength of the radiation field, since the circuit resistance will be much larger than the radiation resistance.

The problem then resolves into relating the measured field strength to the average brightness temperature of that part of the sky covered by the radiation pattern of the system. Even above the reflecting layer the ionosphere will still have an important effect on the fields picked up by the aerial which will be quite complicated. It is to be noted that the gyrofrequency in the high ionosphere lies within the band of frequencies where measurements are required; that is about 10 Mc/s down to the lowest frequency at which measurements can be taken, which is expected to be about 100 kc/s. In this paper a very simple ionosphere is considered where the refractive index is real and lies between 0 and 1. The effects of the earth's magnetic field are ignored, and it is assumed that any gradients in the ionosphere are small so that no appreciable changes occur in the space of one wavelength.

The relation between measured field and brightness temperature will be derived firstly by considering the voltage developed in the radiation resistance of the dipole aerial, and secondly by expressing the energy density in an electromagnetic field in two different ways.

RADIATION RESISTANCE

The radiation resistance of a short electric dipole is given in terms of its effective length, l, by the expression

$$R_r = \frac{Z_o}{6\pi}(kl)^2$$

Z_o = characteristic impedance of free space
k = wave number
l = effective length

If the voltage between the terminals of the open-circuited dipole is V at a frequency f, in a bandwidth B, then the measured field strength

of the radiation field, E_m, is defined by

$$l E_m = V$$

When the aerial is within a black-body enclosure at temperature T, the voltage between its terminals is given by

$\overline{V^2} = 4k\,TB\,R_r$ (the bar denotes a time average)

$$\therefore \quad l^2\,\overline{E_m^2} = 4kTB\,\frac{Z_o}{6\pi}(k\,l)^2 \qquad (1)$$

$$\text{or} \quad T = \frac{\lambda^2}{\frac{8}{3}\pi kB}\,\frac{\overline{E_m^2}}{Z_0}$$

Thus we have related the average brightness temperature of that part of the sky exposed to the radiation pattern of the short dipole aerial with the field strength measured by it. Knowing V and l we can calculate T.

In the calibration of loop antennas it is convenient to use a known C.W. plane wave field. This can be measured directly by means of a field strength meter (i.e. a short electric aerial). In this case the output of the loop antenna and receiver system is measured for a known value of electric field strength perpendicular to the loop axis. When the loop is now immersed in an isotropic radiation field the output of the receiver will record the electric field perpendicular to the loop axis which would give the same output. But, since the average magnetic and electric field strengths in an isotropic radiation field bear the same proportion to each other as those in a C.W. field, measurement of the component of magnetic field in a certain direction will be equivalent to measurement of the electric field in any other given direction. That is, the field strength deduced from the receiver output will be that which would be measured by a short electric dipole, E_m. We may call this field strength the *equivalent C.W. field strength.*

FIELD STRENGTH IN A BLACK-BODY

There is another way of obtaining equation (1) which is more instructive, as it shows how the measured field strength will vary in a simple ionosphere, and also how the radiation resistance depends upon the refractive index of the medium. The energy density in black-body radiation is given by two expressions which are equated thus

$$\frac{8\pi kTB}{c\lambda_2} = \frac{\varepsilon\varepsilon_o\overline{E^2}}{4\pi} \qquad (2)$$

The first is the Raleigh-Jeans approximation and the second is an expression of electromagnetic theory.

Here E is the electric vector of the radiation field,
c is the velocity of electromagnetic waves in the medium,
λ is the wavelength in the medium at the frequency considered.

The electric field along the axis of the dipole aerial, E^1, is related to the actual field by

$$E^1 = E\cos\theta$$

θ being the angle between the electric vector and the dipole axis. The voltage between the aerial terminals is given by

$$V = E^1 \times l$$

and so $E^1 = E_m$

Averaging over all possible θ

$$\overline{E_m^2} = \tfrac{1}{3}\,\overline{E^2}$$

Substitution in (2) then gives as before

$$T = \frac{\lambda^2}{\frac{8}{3}\pi kB}\,\frac{\overline{E_m^2}}{\overline{Z}}$$

where $Z = \dfrac{4\pi}{c\varepsilon\varepsilon_o}$ and is the characteristic impedance of the medium.

This result can be expected from a consideration of what would be absorbed from the radiation field by three mutually perpendicular dipoles.

As before $\quad kTB = \dfrac{\lambda^2}{8\pi}\,\dfrac{E^2}{Z_o}$

where E is now the field strength picked up by an isotropic radiator and kTB is the available power from the radiation field. Since each dipole will take on average one third of the power available for an isotropic radiator $\overline{E_m^2} = \tfrac{1}{3}\,\overline{E^2}$ as before.

REFRACTIVE INDEX OF THE MEDIUM

Rewriting (2) we have $\overline{E^2} = \frac{4\pi}{\varepsilon\varepsilon_0}\frac{8\pi kTB}{c\lambda^2}$.

We consider an isotropic medium with no attenuation for which ε is real and equals $1-X$. The refractive index n is equal to $\varepsilon^{\frac{1}{2}}$.

$$\text{Since} \quad \varepsilon \propto n^2$$
$$c \propto \frac{l}{n}$$
$$\lambda \propto \frac{l}{n},$$

It follows from (2) that $\overline{E^2} \propto n$ or $\sqrt{\overline{E^2}} \propto n^{\frac{1}{2}}$.

Similarly using $\overline{H^2} = \frac{4\pi}{\mu_0}\frac{8\pi kTB}{c\lambda^2}$ and putting $\mu = 1$ we could have obtained the result $\sqrt{\overline{H^2}} \propto n^{3/2}$.

As is to be expected $\frac{\sqrt{\overline{H^2}}}{\sqrt{\overline{E^2}}} \propto n$

since for any electromagnetic wave

$$\frac{H}{E} = \sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}} \propto n.$$

We now define the radiation resistance of a dipole by the formula $\overline{V^2} = 4kTBR_r$ where $\sqrt{V^2}$ is the r.m.s. voltage between the open circuited dipole terminals, and T is the temperature of the black body in which the dipole is placed. This is justified because an aerial placed in an isotropic non-lossy medium will 'see' the temperature of the surrounding enclosure. If the medium were lossy then a proportion of the output from the dipole would correspond with the temperature of the medium, until, as the optical depth of the medium becomes very large, the aerial effectively measures the temperature of the medium alone.

Now if the effective length of the dipole is not appreciably altered by the medium

$$\overline{V^2} \propto \overline{E^2}$$
$$\propto n$$

consequently $R_r \propto n$

Similarly we can define the radiation resistance of a loop antenna by

$$\overline{i^2} = 4kTBG_r$$

where $R_r = \frac{1}{G_r}$ and $\sqrt{i^2}$ is the r.m.s. current picked up by the loop.

But $$R_r^2 \ \overline{i^2} \propto \overline{H^2}$$
$$\propto n^3$$

So for a loop antenna $R_r \propto n^3$.

This result was obtained by Chapman by a consideration of the derivation of the radiation resistance of a loop by Aharoni. He obtained

$$R_r = \frac{\mu_0\omega}{6\pi} A^2 \left(\frac{2\pi}{\lambda_0}\right)^3 \varepsilon^{3/2}$$

It would be desirable to extend these results to cases where the medium is not isotropic and where the medium is lossy.

EXPERIMENTAL RESULTS

Molozzi, Franklin and Tyas[1] have measured the radiation field in the high ionosphere using loop dipoles mounted on satellite 1960η.1. Chapman[2] and Molozzi state that this corresponds to a brightness temperature of $8{\cdot}1 \times 10^6$ °K for the sky at R.A. $5{\cdot}5^h$ Dec. $-30°$. Using the analysis of this paper Smith[3] has found that the brightness temperature of the northern sky is $1{\cdot}3 \times 10^6$ °K and he has used this value to derive the spectral index of the galactic radiation.

REFERENCES

1. A. R. Molozzi, C. A. Franklin and J. P. I. Tyas, *Nature* (*Lond.*) **190**, 616, 1961.
2. J. H. Chapman and A. R. Molozzi, *Nature* (*Lond.*) **191**, 480 (1961).
3. F. G. Smith, *Nature* (*Lond.*) (in course of publication).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 71–76. Printed in Great Britain.

WHISTLER DISPERSIONS FOR A MODEL OF THE IONIC DISTRIBUTION IN THE EXOSPHERE

T. N. L. PATTERSON
Department of Applied Mathematics, The Queen's University of Belfast, Belfast, N. Ireland

(*Received* 25 *August* 1961)

Abstract—Calculations are presented of the dispersions of whistling atmospherics for a model of the ion distributions in the exosphere. Exospheric temperatures of 1000°K, 1250°K, 1500°K and 2000°K are considered at latitudes 30°, 35°, 40° and 45° and corresponding to various transition levels at which n (500 km|O^+) equals n (500 km|H^+).

Comparison is made with two observed whistlers.

INTRODUCTION

In a recent paper Bates and Patterson[1] have discussed the distributions of hydrogen ions and electrons in the Earth's exosphere and have derived analytical expressions for the variations of the number densities with altitude. These expressions enable the dispersion curves of whistlers to be easily calculated.

Some calculations on whistler dispersion have already been reported by Barrington and Nishizaki[2]. Artificial simplifying assumptions were made in these: thus, above a certain transition level the number of protons was equated to the number of electrons while below that level the protons were entirely neglected. This paper describes calculations which are more refined and more extensive.

ADOPTED DISTRIBUTIONS OF ELECTRONS AND PROTONS

Comparison of the distribution of electrons given by the original analytical expression with the distribution recommended by workers at the Naval Research Laboratory (which is based on observational data) shows that the fall-off at great altitudes is not correct (Bates and Patterson[1]). The theoretical distribution through the exosphere must be multiplied by R^{-n}, where R is the distance from the Earth's centre and n is about 3, to bring it into accord with the *NRL* distribution.

To describe the situation at great altitudes realistically it was therefore decided to include an R^{-3} factor* in the proton distribution.

The distributions are then:

$$n(z|H^+) = \frac{(R_e/R)^3\, n(z_e|H^+)\,[n(z_e|e)]^{\frac{1}{2}} \exp \alpha(z, T)}{[(R_e/R)^3 n(z_e|H^+) \exp \alpha(z, T) + n(z_e|O^+) \exp 16\,\alpha(z, T)]^{\frac{1}{2}}} \quad (1)$$

$$n(z|O^+) = \frac{n(z_e|O^+)\,[n(z_e|e)]^{\frac{1}{2}} \exp 16\,\alpha(z, T)}{[(R_e/R)^3 n(z_e|H^+) \exp \alpha(z, T) + n(z_e|O^+) \exp 16\,\alpha(z, T)]^{\frac{1}{2}}} \quad (2)$$

$$n(z|e) = [n(z_e|e)]^{\frac{1}{2}}\,[(R_e/R)^3 n(z_e|H^+) \exp \alpha(z, T) + n(z_e|O^+) \exp 16\,\alpha(z, T)]^{\frac{1}{2}} \quad (3)$$

where:

$$\alpha(z, T) = -\frac{1{\cdot}103}{T}(z-500)\left[\left(1+\frac{z}{R_o}\right)^{-1} - 3{\cdot}73\times 10^{-3}\left(1+\frac{z+500}{2R_o}\right)\right]. \quad (4)$$

R_o is the Earth's radius and T is the exospheric temperature of all the particles. The R^{-3} factor was chosen to become effective at 500 km the altitude of the reference level and the base of the exosphere.

CALCULATIONS

Following Barrington and Nishizaki[2] the

* Contrary to what was earlier believed this factor has no explanation (Johnson[3]).

dispersion of a whistler in a medium containing electrons and protons* may be expressed as:

$$D=\frac{1}{c}\int_0^{\lambda_d}\frac{f_0}{f_H^{\frac{1}{2}}}\left\{\left(1-\frac{f}{f_H}\right)^{-\frac{3}{2}}+\frac{\left[1-\varepsilon\left(\frac{f_{HI}}{f+f_{HI}}\right)^2\right]}{\left[1-\varepsilon\left(\frac{f_{HI}}{f+f_{HI}}\right)\right]^{\frac{1}{2}}}-1\right\} \cdot \left(1+\frac{z}{R_0}\right)(1+4\tan^2\lambda)^{\frac{1}{4}}\,d\lambda, \tag{5}$$

f_0, f_H, f_{HI}, ε and λ_d being respectively the plasma frequency, the electron gyrofrequency, the proton gyrofrequency, the ratio of the number of protons to the number of electrons and the latitude of detection. Implicit in (5) is the assumption of whistler propagation along the geomagnetic line of force from the initiating *sferic*. The geomagnetic field in oersteds is:

$$H=0{\cdot}31\left(1+\frac{z}{R_0}\right)^{-3}\cos\lambda\,(1+4\tan^2\lambda)^{\frac{1}{2}} \tag{6}$$

and the altitude z on the line of force from latitude λ_d at latitude λ is:

$$z=R_0\left[\left(\frac{\cos\lambda}{\cos\lambda_d}\right)^2-1\right]. \tag{7}$$

Further:

$$f_0=8{\cdot}97\times10^3\,[n(z|e)]^{\frac{1}{2}}\ \text{sec.}^{-1}, \tag{8}$$

$$f_H=2{\cdot}79\times10^6\,H\ \text{sec.}^{-1} \tag{9}$$

and:

$$f_{HI}=1{\cdot}52\times10^3\,H\ \text{sec.}^{-1} \tag{10}$$

The electron density and ε can now be expressed as:

$$n(z|e)=n(z_s|O^+)\left[\gamma(z_s)\frac{1{\cdot}25}{R^3}\exp\alpha(z,T)+\exp 16\,\alpha(z,T)\right]^{\frac{1}{2}} \tag{11}$$

and:

$$\varepsilon=\frac{1{\cdot}25\exp\alpha(z,T)}{R^3\left[(1{\cdot}25/R^3)\exp\alpha(z,T)+(1/\gamma(z_s))\exp 16\,\alpha(z,T)\right]} \tag{12}$$

* The direct effect of the oxygen ions is neglected.

with

$$\gamma(z_s)=\frac{n(z_s|H^+)}{n(z_s|O^+)}. \tag{13}$$

R is measured in Earth radii. Combined use of equations (6) to (12) shows that (5) may be expressed in the form:

$$D=[n(z_s|O^+)]^{\frac{1}{2}}\,G(\lambda_d,T,\gamma,f). \tag{14}$$

RESULTS

The function G in (14) has been evaluated by numerical integration for the latitudes 30°, 35°, 40° and 45° and for exospheric temperatures of 1000°K, 1250°K, 1500°K and 2000°K over a frequency range of 0·5 to 9 kc/sec. In each case the values of γ were chosen so as to cause equality of hydrogen ions and oxygen ions at the transition levels 750 km, 1000 km, 1250 km, 1500 km and in some cases 1750 km, determined by

$$\gamma(z_s)=\frac{R^3}{1{\cdot}25}\exp 15\alpha(z_t,T), \tag{15}$$

where z_t is the altitude of the transition level. It was further assumed throughout that

$$n(500|O^+)=4\times10^3\ \text{cm}^{-3}., \tag{16}$$

but, as seen from (14), this only appears as an isolated multiplying factor.

Since (11) and (12) are not valid below the reference level of 500 km, a correction must be applied to all dispersions for the path through the lower ionosphere. This was done by integrating (5) for the region above the 500 km level only and then adding the contribution from a uniform lower ionosphere 400 km in radial thickness, containing 4×10^5 electrons per cm.^{3}

With four possible variables in (14) difficulty arises in attributing an observed dispersion curve to a unique situation. A small vertical shift of the theoretical curve is permissible but the corrections applied to take account of the contribution from the lower ionosphere cannot be seriously in error. In general it is necessary to multiply the theoretical dispersions by a factor K to give agreement with the observed values. This implies that the assumed value

Table 1. Frequencies of minimum dispersions

Latitude	Temperature (°K)	Transition levels (kms)				
		750	1000	1250	1500	1750
		Frequencies in kc/sec				
30°	1000	4·1	2·8	$1·6_5$	—	—
	1250	4·5	3·5	2·5	$1·6_5$	—
	1500	4·7	3·9	$2·8_5$	2·3	—
	2000	—	4·2	$3·4_5$	2·8	2·0
35°	1000	$3·8_5$	$3·0_5$	$2·1_5$	—	—
	1250	$4·1_5$	$3·4_5$	2·7	$2·2_5$	—
	1500	$4·0_5$	$3·7_5$	3·3	$2·6_5$	—
	2000	—	4·0	$3·5_5$	$3·1_5$	$2·6_5$
40°	1000	$3·2_5$	$2·7_5$	2·2	—	—
	1250	$3·3_5$	3·0	$2·7_5$	$2·2_5$	—
	1500	$3·4_5$	3·0	2·8	2·5	—
	2000	—	3·2	$2·9_5$	2·6	2·6
45°	1000	$2·4_5$	2·2	$1·8_5$	—	—
	1250	$2·4_5$	$2·3_5$	$2·2_5$	2·0	—
	1500	$2·4_5$	$2·3_5$	2·2	2·1	—
	2000	—	$2·3_5$	2·3	$2·2_5$	$2·1_5$

Table 2. γ (500) for various transition levels

Transition level (km)	Temperature (°K)			
	1000	1250	1500	2000
750	$2·8\times10^{-2}$	$5·9\times10^{-2}$	$9·6\times10^{-2}$	$1·8\times10^{-1}$
1000	$1·0\times10^{-3}$	$4·2\times10^{-3}$	$1·1\times10^{-2}$	$3·5\times10^{-2}$
1250	$4·5\times10^{-5}$	$3·5\times10^{-4}$	$1·4\times10^{-3}$	$7·8\times10^{-3}$
1500	$2·5\times10^{-6}$	$3·5\times10^{-5}$	$2·1\times10^{-4}$	$1·9\times10^{-3}$
1750	$1·6\times10^{-7}$	$4·1\times10^{-6}$	$3·5\times10^{-5}$	$5·2\times10^{-4}$
2000	$1·3\times10^{-8}$	$5·4\times10^{-7}$	$6·6\times10^{-6}$	$1·5\times10^{-4}$
γ^{BP}(500)	$2·3\times10^{-3}$	$4·0\times10^{-4}$	$7·6\times10^{-5}$	$1·5\times10^{-5}$

of $n(500|O^+)$ should be multiplied by a factor K^2.

However, the location of the frequency at which the minimum dispersion occurs restricts the acceptable values of the parameters. Table 1 gives the minima for all the cases considered. It can be seen that they are quite sensitive to the physical conditions especially at low latitudes. The values of γ (500) corresponding to the temperatures and transition levels considered are given in Table 2, $n(500|O^+)$ being taken to be 4×10^5 cm^{-3} as already noted. The last line of the Table gives the values γ^{BP} (500) which Bates and Patterson[1] estimated the parameter to have. It is not expected that γ^{BP} (500) is in error by a large factor so that some limitation is imposed on the values of γ (500) which are likely. However, since whistler dispersion represents an instantaneous state of a medium which is subject to variations, discrepancies can occur. Table 3 gives the dispersion values calculated for what were considered representative cases. Comparison is made with two whistlers given by Barrington and Nishizaki[2] in Figs. 1 and 2. A satisfactory fit is found for whistler No. 1 with a tempera-

Table 3. Whistler dispersions in sec.$^{\frac{1}{2}}$

f kc sec^{-1}	Transition levels (kms)				Transition levels (kms)			
	750	1000	1250	1500	750	1000	1250	1500
	1000°K/30°				1000°K/35°			
0·5°	49·230	31·927	25·513	—	57·770	33·920	24·605	—
1·0	49·007	31·841	25·495	—	57·294	33·714	24·524	—
2·0	48·474	31·681	25·481	—	56·603	33·469	24·458	—
3·0	48·270	31·656	25·522	—	56·370	33·420	24·478	—
4·0	48·219	31·691	25·588	—	56·330	33·448	24·531	—
5·0	48·246	31·754	25·664	—	56·385	33·515	24·600	—
6·0	48·316	31·835	25·748	—	56·493	33·604	24·679	—
7·0	48·413	31·926	25·836	—	56·634	33·705	24·762	—
8·0	48·529	32·024	25·926	—	56·798	33·815	24·850	—
9·0	48·657	32·127	26·019	—	56·977	33·932	24·940	—
	1000°K/40°				1000°K/45°			
0·5	68·148	37·531	25·358	—	82·676	43·427	27·511	—
1·0	67·469	37·231	25·229	—	81·892	43·068	27·354	—
2·0	66·726	36·955	25·142	—	81·267	42·833	27·281	—
3·0	66·539	36·919	25·162	—	81·280	42·878	27·334	—
4·0	66·575	36·975	25·222	—	81·541	43·028	27·434	—
5·0	66·719	37·075	25·293	—	81·925	43·229	27·555	—
6·0	66·924	37·200	25·391	—	82·382	43·460	27·690	—
7·0	67·168	37·341	25·499	—	82·886	43·711	27·833	—
8·0	67·438	37·492	25·589	—	83·426	43·976	27·983	—
9·0	67·727	37·651	25·694	—	83·993	44·253	28·138	—
	1250°K/30°				1250°K/35°			
0·5	58·201	39·395	31·157	27·491	70·980	44·517	32·296	26·554
1·0	57·913	39·246	31·094	27·478	70·349	44·179	32·121	26·473
2·0	57·211	38·957	30·997	27·473	69·435	43·763	31·941	26·411
3·0	56·921	38·868	30·997	27·515	69·109	43·644	31·914	26·429
4·0	56·825	38·868	31·038	27·576	69·032	43·645	31·948	26·478
5·0	56·825	38·912	31·100	27·646	69·076	43·703	32·011	26·541
6·0	56·880	38·980	31·174	27·722	69·189	43·795	32·090	26·612
7·0	56·969	39·064	31·255	27·801	69·345	43·906	32·178	26·688
8·0	57·082	39·159	31·340	27·882	69·529	44·031	32·274	26·768
9·0	57·210	39·261	31·430	27·966	69·733	44·165	32·374	26·850
	1250°K/40°				1250°K/45°			
0·5	86·023	51·445	35·164	27·162	107·673	62·079	40·289	29·429
1·0	85·103	50·951	34·900	27·022	106·574	61·479	39·965	29·251
2·0	84·110	50·485	34·685	26·932	105·732	61·080	39·782	29·174
3·0	83·849	50·393	34·668	26·951	105·753	61·128	39·838	29·229
4·0	83·880	50·445	34·725	27·008	106·106	61·343	39·979	29·331
5·0	84·051	50·565	34·817	27·083	106·623	61·638	40·162	29·455
6·0	84·302	50·723	34·928	27·169	107·237	61·981	40·369	29·592
7·0	84·604	50·906	35·051	27·261	107·915	62·355	40·593	29·738
8·0	84·940	51·106	35·183	27·358	108·639	62·752	40·829	29·890
9·0	85·301	51·317	35·321	27·458	109·401	63·168	41·074	30·048
	1500°K/30°				1500°K/35°			
0·5	65·440	46·300	36·878	32·204	81·740	54·414	40·267	32·687
1·0	65·110	46·102	36·775	32·165	80·990	53·958	39·995	32·536
2·0	64·281	45·700	36·599	32·112	79·900	53·385	39·699	32·396
3·0	63·928	45·557	36·562	32·131	79·503	53·206	39·629	32·385

Table 3 (*continued*)

f kc sec^{-1}	Transition levels (kms)				Transition levels (kms)			
	750	1000	1250	1500	750	1000	1250	1500
4·0	63·799	45·531	36·587	32·182	79·400	53·185	39·649	32·425
5·0	63·783	45·562	36·641	32·248	79·441	53·242	39·712	32·488
6·0	63·831	45·626	36·712	32·323	79·562	53·342	39·798	32·565
7·0	63·919	45·710	36·793	32·403	79·733	53·469	39·899	32·650
8·0	64·033	45·808	36·882	32·487	79·938	53·613	40·009	32·740
9·0	64·166	45·915	36·975	32·573	80·167	53·770	40·127	32·834
	1500°K/40°				1500°K/45°			
0·5	101·020	64·837	45·713	35·247	129·047	80·531	54·531	40·100
1·0	99·903	64·160	45·302	34·997	127·682	79·698	54·023	39·784
2·0	98·710	63·561	44·954	34·814	126·667	79·146	53·725	39·627
3·0	98·394	63·377	44·904	34·810	126·704	79·206	53·791	39·692
4·0	98·425	63·430	44·964	34·871	127·143	79·492	53·988	39·835
5·0	98·624	63·575	45·075	34·962	127·781	79·889	54·248	40·016
6·0	98·919	63·772	45·216	35·070	128·535	80·349	54·545	40·218
7·0	99·273	63·999	45·376	35·188	129·368	80·853	54·866	40·435
8·0	99·669	64·253	45·547	35·314	130·257	81·389	55·206	40·663
9·0	100·095	64·521	45·728	35·446	131·192	81·949	55·561	40·900

f kc sec^{-1}	Transition levels (kms)				Transition levels (kms)			
	1000	1250	1500	1750	1000	1250	1500	1750
	2000°K/30°				2000°K/35°			
0·5	57·901	47·465	41·559	38·308	71·375	55·420	45·694	39·543
1·0	57·644	47·306	41·481	38·293	70·723	54·975	45·392	39·353
2·0	57·066	46·993	41·336	38·257	69·900	54·462	45·081	39·179
3·0	56·841	46·895	41·317	38·290	69·624	54·316	45·015	39·164
4·0	56·776	46·893	41·355	38·352	69·572	54·314	45·043	39·209
5·0	56·791	46·939	41·419	38·429	69·631	54·382	45·115	39·282
6·0	56·853	47·012	41·500	38·515	69·751	54·488	45·212	39·371
7·0	56·942	47·101	41·589	38·606	69·908	54·619	45·324	39·470
8·0	57·051	47·202	41·686	38·701	70·092	54·765	45·447	39·575
9·0	57·173	47·310	41·787	38·798	70·293	54·923	45·577	39·685
	2000°K/40°				2000°K/45°			
0·5	88·195	66·269	52·582	43·851	113·621	83·213	63·963	51·485
1·0	87·211	65·580	52·098	43·512	112·381	82·341	63·343	51·043
2·0	86·268	64·979	51·716	43·272	111·574	81·826	63·015	50·836
3·0	86·053	64·869	51·669	43·265	111·666	81·923	63·111	50·928
4·0	86·114	64·941	51·744	43·344	112·090	82·241	63·356	51·124
5·0	86·309	65·099	51·877	43·454	112·675	82·665	63·673	51·368
6·0	86·578	65·305	52·041	43·590	113·356	83·151	64·032	51·642
7·0	86·895	65·541	52·225	43·739	114·100	83·679	64·420	51·935
8·0	87·244	65·799	52·423	43·897	114·891	84·238	64·828	52·243
9·0	87·617	66·072	52·631	44·062	115·719	84·822	65·253	52·563

ture of 1250°K and a 750 km transition level, corresponding to:

$$n(500|O^+) = 1{\cdot}4 \times 10^5 \text{ cm}^{-3}$$

and

$$n(500|H^+) = 8{\cdot}5 \times 10^3 \text{ cm}^{-3}.$$

With whistler No. 2 the fit is better with a temperature of 1000°K and a 750 km transition level, corresponding to:

$$n(500|O^+) = 3{\cdot}0 \times 10^5 \text{ cm}^{-3}$$

and

$$n(500|H^+) = 8{\cdot}3 \times 10^3 \text{ cm}^{-3}.$$

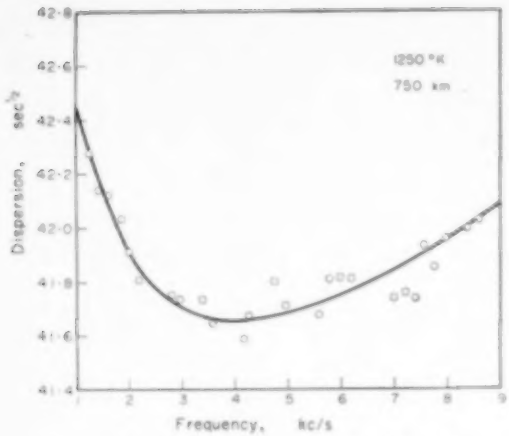


Fig. 1. Dispersion of whistler No. 1 at Wakkanai, latitude 35·3° on 9 March 1958 at 2–5 hrs. J.S.T. The circles denote the observed values.

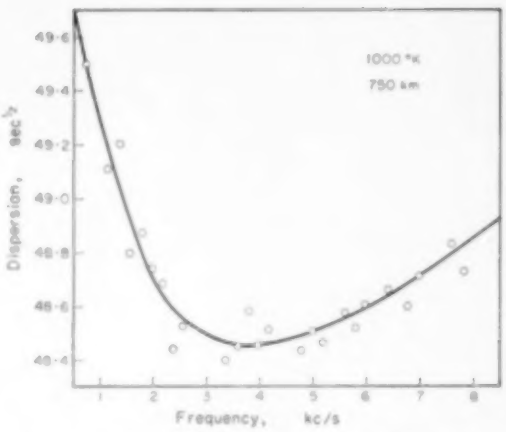


Fig. 2. Dispersion of whistler No. 2 at Wakkanai, latitude 35·3° on 24 December 1957 at 1600 hrs. J.S.T. The circles denote the observed values.

Acknowledgments—The author thanks Professor D. R. Bates for his interest in and suggestions on this work and for his many helpful discussions.

The research has been supported by the Air Force Cambridge Research Center, Geophysics Research Directorate of the Air Research and Development Command, U.S. Air Force, through its European Office, under grant AF–EOARDC 61-16.

REFERENCES

1. D. R. Bates and T. N. L. Patterson, *Plan. Space Sci.* **5**, 328 (1961).
2. R. E. Barrington and I. Nishizaki, *Canad. J. Phys.*, **38**, 1642 (1960).
3. F. S. Johnson, *Agardograph*, **42**, 19 (1959).

SELECTED PAPERS AND ABSTRACTS TRANSLATED FROM

ARTIFICIAL EARTH SATELLITES
U.S.S.R.

(ISKUSSTVENNYE SPUTNIKI ZEMLI)

(PUBLISHED BY THE U.S.S.R. ACADEMY OF SCIENCES,
MOSCOW)

Translations of any papers abstracted here may be obtained from Pergamon Institute at a reasonable cost. Enquiries should be addressed to the Administrative Secretary, Pergamon Institute, at either Headington Hill Hall, Oxford, England, or 122 East 55th Street, New York 22, N.Y., U.S.A.

Any papers marked * have been selected for publication in this section of future issues of *Planetary and Space Science*.

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 79–81. Printed in Great Britain.

THE EQUATIONS OF DISTURBED MOTION IN THE KEPLER PROBLEM

A. I. LUR'E

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 4, p. 82 (1960)

The equations for disturbed motion of a planet were already known in part by Newton. The history of the problem and the results of the equations were set out in Tisserand's *Traité de Mécanique Céleste*[1] and by A. N. Krylov[2].

Tisserand, following the general methods of the theory of disturbed motion, calculated the Lagrange factors for the elliptical elements of the orbit; Krylov's conclusions were based on geometrical constructions. The results of these equations are set out in the work[3]. The conclusion here proposed is based on the direct application of the method of variation of arbitrary constants. The equation of the elliptical orbit is written in vectorial form

$$\mathbf{r}=\frac{a(1-e^2)}{1+e\cos\phi}\mathbf{e}_r=r\mathbf{e}_r \tag{1}$$

where $\mathbf{e}_r$ is the unit vector from the centre of gravity to the travelling point; a, e are respectively the large semi-axis and eccentricity of orbit, $\cos\phi=\mathbf{e}_r\cdot\mathbf{e}_1$, where $\mathbf{e}_1$ is the unit vector of direction at perigee (along the large semi-axis of the orbit).

We introduce into the study the orthogonal trihedral of the unit vectors, $\mathbf{e}_r$, $\mathbf{e}_\phi$, $\mathbf{e}_3=(\mathbf{e}_r\mathbf{e}_\phi)$; $\mathbf{e}_\phi$ is the unit vector directed towards increasing angle ϕ in the orbital plane perpendicular to $\mathbf{e}_r$; $\mathbf{e}_3$ determines the orbital plane in the undisturbed path. In the undisturbed path this trihedral has an angular velocity $\dot\phi\mathbf{e}_3$, so that

$$\dot{\mathbf{e}}_r=\dot\phi\mathbf{e}_\phi,\dot{\mathbf{e}}_\phi=-\dot\phi\mathbf{e}_r,\dot{\mathbf{e}}_3=0 \tag{2}$$

and, from the theorem of surface,

$$\dot\phi=\frac{\sqrt{\mu a(1-e^2)}}{r^2} \tag{3}$$

where μ is the proportionality coefficient of the law of gravitation.

The position of the orbital plane is governed by the longitude of the ascending node Ω yielding the direction of the unit vector of $\mathbf{n}$ node lines, and the angle of inclination i of the orbital plane to the plane $0\xi\eta$ of the system of stationary axes $0\xi\eta\zeta$: the position of the perigee in the orbital plane is given by the angular separation of the perigee from the node ω so that $\cos\omega=\mathbf{n}\cdot\mathbf{e}_1$.

It follows from (1), (2), (3) that the velocity vector, in undisturbed motion, is equal to

$$\mathbf{v}=\dot{\mathbf{r}}=\sqrt{\frac{\mu}{a}}\cdot\frac{1}{\sqrt{1-e^2}}[e\sin\phi\mathbf{e}_r+(1+e\cos\phi)\,\mathbf{e}_\phi] \tag{4}$$

whilst the acceleration vector

$$\mathbf{w}=\dot{\mathbf{v}}=-\frac{\mu}{r^2}\mathbf{e}_r \tag{5}$$

Following the method of variation of arbitrary constants, we will maintain for the vectors $\mathbf{r}$ and $\mathbf{v}$ the same expressions (1) and (4) in disturbed motion, as in undisturbed motion; however, the elliptical elements of the orbit a, e, Ω, i, ω will not now be constants but unknown functions of time. The angular velocity $\boldsymbol{\omega}$ of the trihedral $\mathbf{e}_r$, $\mathbf{e}_\phi$, $\mathbf{e}_3$ in disturbed motion is equal to

$$\boldsymbol{\omega}=\mathbf{k}\dot\Omega+\mathbf{n}\frac{di}{dt}+\mathbf{e}_3(\dot\omega+\dot\phi) \tag{6}$$

where $\mathbf{k}$ is the unit vector on the axis $O\zeta$. The projections along the axis of the trihedral $\mathbf{e}_r$, $\mathbf{e}\phi$, $\mathbf{e}_3$ are determined by the known formulae

$$\begin{aligned}\omega_r&=\dot\Omega\sin i\sin u+\frac{di}{dt}\cos u\\ \omega_\phi&=\dot\Omega\sin i\cos u-\frac{di}{dt}\sin u\\ \omega_3&=\dot\Omega\cos i+\dot\omega+\dot\phi=\omega_3'+\dot\phi\end{aligned} \tag{7}$$

where $u=\omega+\phi$. It is observed again that $\dot\phi$ in

these formulae of disturbed motion differs from the value determined from (3); we will represent the latter as $\dot{\phi}^0$ (the zero superscript will be used to differentiate the value in undisturbed motion).

From the formulae of unit vector differentiation we have

$$\dot{\mathbf{e}}_r = [\boldsymbol{\omega}\mathbf{e}_r] = -\omega_\phi \mathbf{e}_3 + (\omega_3' + \dot{\phi})\,\mathbf{e}_\phi$$
$$\dot{\mathbf{e}}_\phi = [\boldsymbol{\omega}\mathbf{e}_\phi] = \omega_r \mathbf{e}_3 - (\omega_3' + \dot{\phi})\,\mathbf{e}_r$$
$$\dot{\mathbf{e}}_3 = [\boldsymbol{\omega}\mathbf{e}_3] = -\omega_r \mathbf{e}_\phi + \omega_\phi \mathbf{e}_r \qquad (8)$$

Making use now of the equalities

$$\dot{\mathbf{r}} = \mathbf{v} = \mathbf{v}^0, \quad \dot{\mathbf{v}} = \mathbf{w}^0 + \mathbf{F} \qquad (9)$$

where $\mathbf{F}$ is the additional force acting on the point in disturbed motion, we obtain, after carrying out differentiation, taking into account the relations (8), the equations

$$\mathbf{v} = \dot{\mathbf{r}} = \mathbf{e}_r \left(\frac{\partial r}{\partial \phi}\dot{\phi} + \frac{\partial r}{\partial a}\dot{a} + \frac{\partial r}{\partial e}\dot{e}\right) +$$
$$+ r[(\omega_3' + \dot{\phi})\mathbf{e}_\phi - \omega_\phi \mathbf{e}_3] = \left(\mathbf{e}_r \frac{\partial r}{\partial \phi} + \mathbf{e}_\phi r\right)\dot{\phi}^0 =$$
$$= \sqrt{\frac{\mu}{a}}\frac{1}{\sqrt{1-e^2}}[e\sin\phi\,\mathbf{e}_r + (1+e\cos\phi)\,\mathbf{e}_\phi] =$$
$$= v_r \mathbf{e}_r + v_\phi \mathbf{e}_\phi \qquad (10)$$

$$\dot{\mathbf{v}} = \left(\frac{\partial v_r}{\partial a}\dot{a} + \frac{\partial v_r}{\partial e}\dot{e} + \frac{\partial v_r}{\partial \phi}\dot{\phi}\right)\mathbf{e}_r +$$
$$+ \left(\frac{\partial v_\phi}{\partial a}\dot{a} + \frac{\partial v_\phi}{\partial e}\dot{e} + \frac{\partial v_\phi}{\partial \phi}\dot{\phi}\right)\mathbf{e}_\phi + v_r[-\omega_\phi \mathbf{e}_3 +$$
$$+ (\omega_3' + \dot{\phi})\,\mathbf{e}_\phi] + v_\phi[\omega_r \mathbf{e}_3 - (\omega_3' + \dot{\phi})\,\mathbf{e}_r] = -\frac{\mu}{r^2}\mathbf{e}_r + \mathbf{F} \qquad (11)$$

From (10) we obtain three equations

$$\omega_\phi = 0; \quad \omega_3' + \dot{\phi} = \dot{\phi}^0; \quad -\frac{\partial r}{\partial \phi}\omega_3' + \frac{\partial r}{\partial \phi}\dot{a} + \frac{\partial r}{\partial e}\dot{e} = 0 \qquad (12)$$

The last of these equations in the developed form will be

$$\omega_3' e\sin\phi - \frac{\dot{a}}{a}(1+e\cos\phi) +$$
$$+ \frac{2e + e^2\cos\phi + \cos\phi}{1-e^2}\dot{e} = 0 \qquad (13)$$

Employing the relations (12), the equations obtained from the vector equalities (11) can be written in the form

$$-\frac{\dot{a}}{2a}e\sin\phi + \frac{\dot{e}}{1-e^2}\sin\phi - \omega_3' e\cos\phi =$$
$$= \sqrt{\frac{a}{\mu}}\sqrt{1-e^2}F_r,$$
$$-\frac{\dot{a}}{2a}(1+e\cos\phi) + \frac{\dot{e}}{1-e^2}(\cos\phi + e) + \omega_3' e\sin\phi =$$
$$= \sqrt{\frac{a}{\mu}}\sqrt{1-e^2}F_\phi$$
$$\omega_r = \sqrt{\frac{a}{\mu}}\frac{\sqrt{1-e^2}}{1+e\cos\phi}F_3 \qquad (14)$$

From the first equation (12) and the last (14), writing in the values ω_r and ω_ϕ from (7), we find the equations of disturbed motion for the elements Ω and i

$$\frac{di}{dt} = \sqrt{\frac{a}{\mu}}\frac{\sqrt{1-e^2}}{1+e\cos\phi}F_3\cos u; \quad \dot{\Omega}\sin i =$$
$$= \sqrt{\frac{a}{\mu}}\frac{\sqrt{1-e^2}}{1+e\cos\phi}F_3\sin u \qquad (15)$$

From (13) and (14) we obtain

$$\dot{e} = \sqrt{\frac{a}{\mu}}\sqrt{1-e^2}\left(F_r\sin\phi + \right.$$
$$\left. + \frac{e + 2\cos\phi + e\cos^2\phi}{1+e\cos\phi}F_\phi\right)\frac{\dot{a}}{2a} =$$
$$= \sqrt{\frac{a}{\mu}}\frac{1}{\sqrt{1-e^2}}[F_r e\sin\phi + (1+e\cos\phi)F_\phi] \quad \omega_3' =$$
$$\sqrt{\frac{a}{\mu}}\cdot\frac{\sqrt{1-e^2}}{e}\left(-F_r\cos\phi + \frac{2+e\cos\phi}{1+e\cos\phi}F_\phi\sin\phi\right) =$$
$$= \dot{\Omega}\cos i + \dot{\omega} \qquad (16)$$

Finally the second equation (12) yields the relation governing the disturbance for the time t_0 of passage through perigee. If we call

$$\zeta = n(t - t_0) = \frac{\sqrt{\mu}}{a\sqrt{a}}(t - t_0)$$

then the relation is written in the form

$$\frac{\dot{\zeta} - n}{\sqrt{1-e^2}} + \omega_3' = -2\sqrt{\frac{a}{\mu}}\frac{\sqrt{1-e^2}}{1+e\cos\phi}F_r \qquad (17)$$

Hence the result is based on application of

the method of variation of arbitrary constants to the expressions for the radius-vector of the planet and its speed vector, using the well-known formulae for differentiation of unit vectors of the moving trihedral of axes. This result represents the most economical and simple of all those available.

REFERENCES

1. F. TISSERAND, *Traité de mécanique céleste*. Vol. 1, Chapter 10 (1889).
2. A. N. KRYLOV, Sur la variation des éléments des orbites elliptiques de planètes. Collected works, Vol. 6 (1915).
3. G. N. DUBOSHIN, *Introduction to celestial mechanics*. Gostekhizdat, Moscow (1938). (In Russian).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 82–85. Printed in Great Britain.

INVESTIGATION OF METEORIC PARTICLES ON THE THIRD SOVIET ARTIFICIAL EARTH SATELLITE

T. N. NAZAROVA

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 4, p. 165 (1960)

In the solar system, in addition to the planets and their satellites, asteroids and comets there is a multiplicity of smaller bodies collectively termed meteoric matter.

Up to the present time the main information regarding meteoric bodies entering the Earth's atmosphere from interplanetary space has been obtained by methods of astronomy.

Visual, photographic and radar tracking of meteors, the observation of Zodiacal light and the Fraunhofer component of the solar corona, in addition to various methods for studying meteoric matter reaching the Earth's surface, led to the establishment of a physical theory of meteors and are used for determining the space density of meteoric bodies in the solar system[1–7]. The first three methods allow determination of the number of particles entering the Earth's atmosphere, together with their masses and size distribution. However, the possibilities afforded by these methods are limited by the possibility at present of recording particles with masses from 10^{-4}g and over. In order to obtain information about finer particles use can be made of the integral characteristics of meteoric matter, obtained from observations of Zodiacal light and having recourse to extrapolation.

As seen from Table 1, photometric investigations yield sharply different values of the space density of the dust matter in the vicinity of the Earth's orbit.

The study of meteoric matter—its cosmic nature (number, spatial density, the mass of meteoric bodies, etc.) location and function in the solar system—is important to astronomy, geophysics and for the solution of problems of the evolution and origin of planetary systems, since it affords clarification of a number of questions confronting present-day cosmogonic hypotheses.

Table 1.

Author	Density, ρ, g/cm^3
V. G. Fesenkov (1947)	6×10^{-23}
Allen (1947)	4×10^{-23}
Van de Holst (1947)	3×10^{-21}
Behr and Siedentopf (1953)	10^{-23}
Alsässer (1954)	2×10^{-23}
Minaart (1955)	6×10^{-22}
Siedentopf (1955)	$2\text{–}4\times10^{-22}$

A work[8] has recently been completed which examines new viewpoints regarding the influence of micrometeoric particles on the geophysical processes in the atmosphere.

The investigation of meteoric matter is also necessary for the solution of a number of practical problems, in particular problems connected with the study of the conditions governing the motion of rockets and artificial earth satellites in cosmic space[9–11].

It is known that the danger of space ships meeting a meteoric body capable of destroying it or destroying the hermetic sealing of the cabin is slight. The space density of meteoric bodies capable of severely damaging a rocket has been closely studied. Such bodies, on entering the Earth's atmosphere, generate light or ionization or fall to the Earth's surface in the form of meteorites. By analyzing the data obtained from observations of meteors on meeting the Earth it is possible to obtain information regarding the space density, the mass of meteoric bodies and an entire range of further important information.

The prolonged action of fine meteoric particles on the shell of the rocket or satellite produces gradual erosion of the surface, and also deterioration of the solar batteries and optical system. Hence the study of micrometeoric particles is of great scientific and

purely practical importance, and is therefore a matter of immediate interest, since, as indicated above, in contrast to particles with masses of 10^{-4}g and over, when dealing with particles with smaller masses there are only the integral characteristics or data obtained by extrapolation.

The data characterizing an individual particle can obviously be obtained only by direct methods employing rockets and in particular employing artificial earth satellites continuing over a prolonged period in the outer atmosphere.

When a meteoric body comes into collision with an obstacle a number of physical processes occur, of which use can be made in the design of a recording instrument intended for installing on rockets and satellites. These instruments take the form of special crystal microphones, photomultipliers, recording the illumination occurring with the impact, pickups, consisting of a number of thin films which are lacerated on impact with the particle, pickups, working on the accelerometer principle, and a variant of the piezoelectric pickup. The surface erosion of the satellite owing to impact with a micrometeoric particle is determined by the variation in the resistance of a thin metal layer and also by the variation in the intensity of the β-emission[12–18].

On the third Soviet artificial earth satellite use was made of an instrument for recording the number of particle impacts and the number of impulses of the pickup material, liberated with explosion of the meteoric particle on its surface[19]. A number of hypotheses exist regarding the connection of the impulse "discharge" with the particle mass, its speed, diameter, etc.

A theoretical calculation carried out by K. P. Stanyukovich[20] has shown that for high speeds the recorded impulse is proportional to the energy of the striking particle.

$$I \approx \bar{\theta} \sqrt{\frac{2\alpha}{\varepsilon_k}} E_0$$

where $\bar{\theta}$ is a coefficient depending on the instability of the process and particularly on the angular distribution of the ejected mass (for a typical explosion of a meteoric body $\bar{\theta} \approx \frac{1}{3}$), ε_k is the energy density of the crystal lattice of the vaporized body, whilst the value of the dimensionless coefficient α is a function of the properties of the vaporized medium. Hence

$$I \sim F(E_0).$$

Now according to M. A. Lavrent'ev[21] [see correction on p. 205 of the present issue, *Iskusstvennye Sputniki Zemli*, No. 4, 1960]

$$I = A a^3 v_0^{1.6} T_0^{-0.3}$$

where A is a constant quantity, a^3 is the mass of the body, v_0 is its speed, T_0 is the minimum density of the body necessary for converting the substance to the gaseous state.

Starting from the simplest theoretical law relating the impulse received by the pickup to the particle energy, assuming at the same time that the mean particle velocity is 40 km/s, then the masses of the recorded particles can be determined from the relation $I = AE_0$. Since the Fifth Assembly of the International Geophysical Year, the proportionality coefficient A between the discharge impulse and the kinetic energy of the particle E_0 has been defined more precisely. A variation in the coefficient A of course would result in a change in determination of the mass of the recorded particle. At present the coefficient A is determined with an accuracy in the order of a half.

Measurement of the "discharge" impulses on the third satellite was carried out by means of a ballistic piezoelectric pickup[22]. Electrical signals occurring on the pickup as the result of impact of the meteoric particle were separated according to amplitude into four channels by the amplifier-converter, counting the number of impulses in each channel.

The instrument was calibrated for recording particles with masses within the ranges $8 \times 10^{-9} - 2{\cdot}65 \times 10^{-8}$g, $2{\cdot}65 \times 10^{-8} - 1{\cdot}5 \times 10^{-7}$g, $1{\cdot}5 \times 10^{-7} - 5{\cdot}6 \times 10^{-6}$g and greater than $5{\cdot}6 \times 10^{-6}$g. A signal was passed to telemetering after 32 impacts had been stored in the smallest particle mass range, after storing 16 and 4 in the two next ranges and after each impact in the last range.

The instrument did not record vibration and noise originating from within the satellite, since

its thresholds were set above the level of this interference.

The piezoelement installed on the satellite recorded not only particles striking its surface. It also possessed a certain sensitivity to impacts on the shell of the satellite and this was taken into account in working the experimental data.

Four piezoelements with an overall area of 3410 cm^2 (including the element housing) were installed for recording meteoric particles on the bottom of the satellite.

Starting, as indicated above, from the theoretical law relating the impulse received by the pickup to the energy of the meteoric particle and assuming that the mean particle velocity is 40 km/s, then during the experiment records were obtained of impact with the pickups of particles with masses from 8×10^{-9} to $2{\cdot}65 \times 10^{-8}$g possessing an energy in the order of 10^4–10^5 ergs.

On 15 May, during the recording sessions of this parameter on the telemetering equipment, the frequency of impacts was 4 to 11 per m^2 per sec, and on 16 and 17 May it fell to 4×10^{-3} impacts . m^{-2} . s^{-1}, subsequently reaching $5{\cdot}3 \times 10^{-4}$ impacts . m^{-2} . s^{-1}. During the next 8 days not one actuation was recorded, that is to say the amplifier-converter did not accumulate the necessary number of impacts for passing to the telemetering equipment; the intensity of the meteoric particle flux was less than 10^{-4}g per m^2 per sec.

We will set out some of the features observed on 15 May: —

1. Despite the wide dispersion of the points (particularly at apogee) minima can be observed with a period in the order of 150 sec (Fig. 1).

Table 2.

Loop number	Altitude, km	No. of impacts, $m^{-2} \times sec^{-1}$
2	1300–1500	5
2	1700–1880	11
3	400–600	7
5	400–700	4

The minimum number of impacts is recorded by the pickup when the satellite is turned bottom facing the Earth. During this period the cone is functioning mainly and the bottom

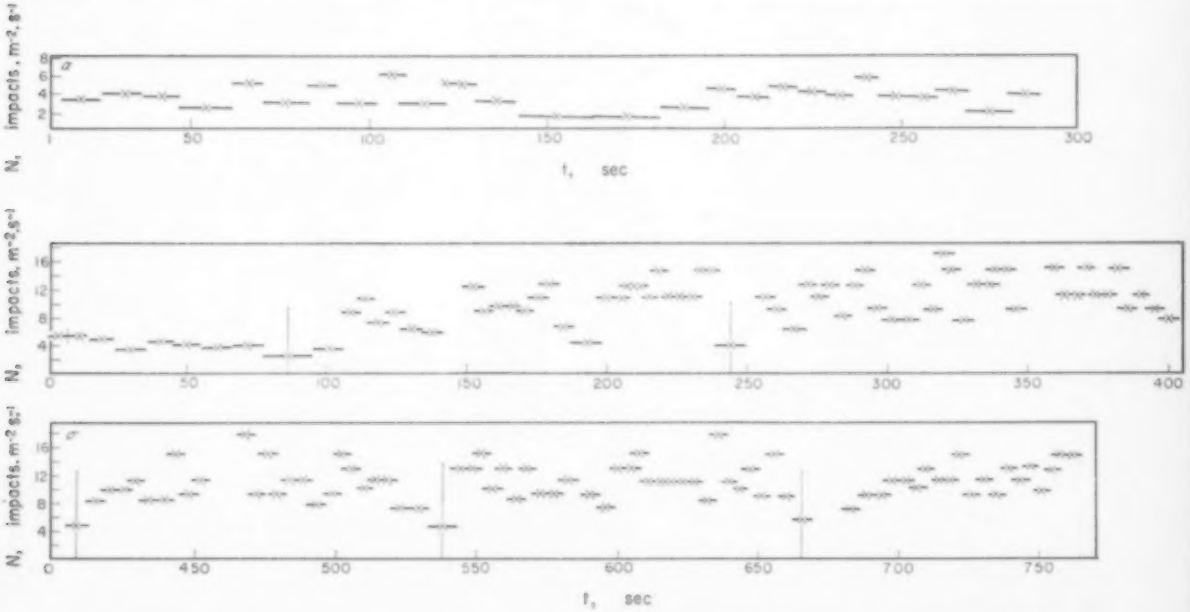


Fig. 1. Number of impacts recorded by the pickups at altitudes: a—1300–1500 km; b—1700–1880 km.

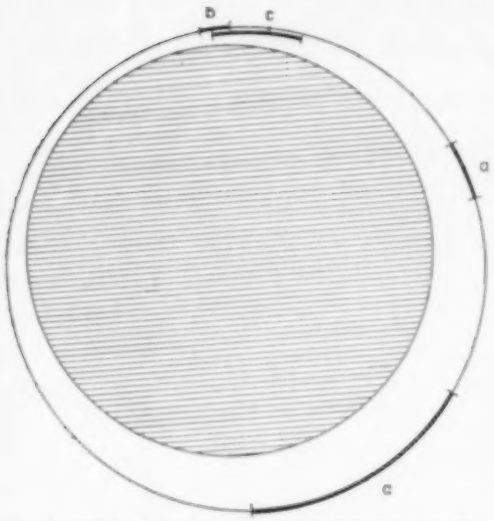


Fig. 2. Location of the sections where meteoric particle impacts were recorded along the satellite orbit. a—second loop; b—third loop; c—fifth loop.

only partly (by way of the corrugation).

2. The frequency of impacts varies with time as the satellite travels along the orbit (Table 2).

The position of the recorded sections along the orbit are shown in Fig. 2. The diagram shows that the number of impacts varies with variation in the position of the satellite along the orbit, but not with variation in altitude.

3. A sharp increase in frequency of impacts is observed only on 15 May. In the succeeding days the frequency of impacts was one-third to one-quarter.

It is difficult at present to say what exactly was the cause of the high impact frequency. This cannot be confidently ascribed to the fact that the satellite intersected a meteor swarm at that time (owing to excessive impact frequency[23]) although it is known from the observation data that the space density of meteoric bodies in the Draconid swarm is thousands of times greater than the density of active drift swarms where it in turn fluctuates over the range of two orders of magnitude. Subsequent investigations will bring greater clarity to this problem.

In conclusion the author expresses his thanks to O. D. Komissarov, L. N. Neugodov, A. A. Trukhachev, L. Z. Rusakov, A. K. Bektabegov and G. M. Kurtev for active participation in the work.

REFERENCES

1. P. G. WATSON, *Between the planets.* Harvard University Press, Cambridge, Mass. (1956).
2. F. L. WHIPPLE, *Physics and Medicine of the Upper Atmosphere.* The University of New Mexico Press, New Mexico (1958).
3. B. YU. LEVIN, *Physical theory of meteors and meteoric matter in the solar system.* (In Russian.) Moscow. Acad. Sci. U.S.S.R. (1956).
4. V. G. FESENKOV, *Meteoric matter in interplanetary space.* (In Russian.) Moscow. Acad. Sci. U.S.S.R. (1947).
5. A. C. B. LOVELL, *Meteor astronomy.* (In Russian.) Fizmatgiz, Moscow (1958). Clarendon Press, Oxford (1954).
6. I. S. ASTAPOVICH, *Meteoric phenomena in the Earth's atmosphere.* (In Russian.) Fizmatgiz, Moscow (1958).
7. L. A. KATASEV, *Photographic methods of meteor astronomy.* (In Russian.) Gostekhizdat, Moscow (1957).
8. B. A. MIRTOV, *Isk. Sput. Zemli*, **4**, 118 (1960).
9. M. W. OVENDEN, *J. Brit. Interplan. Soc.*, **10**, No. 6 (1951).
10. G. GRIMMINGER, *J. Appl. Phys.*, **19**, 947 (1948).
11. F. L. WHIPPLE, *Vistas in Aeronautics*, Pergamon Press, New York (1958).
12. M. DUBIN, *Rocket exploration of the atmosphere.* Pergamon Press, London (1954).
13. O. E. BERG and L. H. MEREDITH, *J. Geophys. Res.* **61**, No. 4 (1956).
14. E. MANRING and M. DUBIN, *Some preliminary reports of experiments in satellites 1958 alpha and 1958 gamma.* IGY World Data Center A, Rockets and Satellites, National Academy of Sciences, Washington, D.C.
15. M. DUBIN, *Cosmic débris of interplanetary space.* Proc. of the Second OSR Astronautics Conference. Denver, Colorado (1958).
16. H. E. LAGOW, D. H. SCHAEFER and I. C. SCHAFFERT, *Micrometeorite impact measurements on a 20″ diameter sphere at 700 to 2500 kilometers altitude*, 5th Assembly International Geophysical Year Specialist Committee (1958).
17. S. F. SINGER, *Jet propulsion*, **26**, No. 12 (1956).
18. E. R. MANRING, *Planet. Space Sci.*, **1**, 27 (1959).
19. M. A. ISAKOVICH, N. A. ROI, *Isk. Sput. Zemli.*, **2**, 81 (1958).
20. S. A. BAUM, S. A. KAPLAN, K. P. STANYUKOVICH. *Introduction to cosmic gas dynamics.* (In Russian.) Fizmatizdat, Moscow (1958).
21. M. A. LAVRENT'YEV, *Isk. Sput. Zemli*, **3**, 61 (1959). *Ibid.*, **4**, 205 (1960).
22. O. D. KOMISSAROV, T. N. NAZAROVA *et al.*, *Isk. Sput. Zemli.*, **2**, 54 (1958).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 86–96. Printed in Great Britain.

THE PATH OF AN ARTIFICIAL EARTH SATELLITE IN THE NORMAL GRAVITATIONAL FIELD OF THE EARTH

M. D. KISLIK

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 4, p. 3 (1960)

The present paper examines the path of an artificial earth satellite in the normal gravitational field of the earth disregarding air resistance and gravity anomalies. The results can be applied to the calculation of satellite orbits at high altitudes, and to the qualitative analysis of the influence of the earth's oblateness on the satellite path.

The potential V of the normal terrestrial gravitational field (the potential of the gravitational field of the total earth ellipsoid[1, 2]) at a point with co-ordinates r and ϕ can be presented in the form of a series[3]:

$$V = \frac{b_0}{r} - \frac{b_2}{r^3} P_2(\sin\phi) + \frac{b_4}{r^5} P_4(\sin\phi) - \ldots, \tag{1}$$

where b_0, b_2, $b_4 \ldots$ are constants, $P_2(\sin\phi)$, $P_4(\sin\phi) \ldots$ are Legendre polynomials of the 2, 4 ... order, r is the distance to the centre of the earth's mass and ϕ is the geocentric latitude.

In the expansion of (1) the terms containing Legendre polynomials of the sixth and higher orders can be disregarded. The forces corresponding to these terms, throughout the entire space surrounding the earth, comprise fractions of a milligal[1]. The term $\frac{b_4}{r^5} P_4(\sin\phi)$ gives rise to an acceleration of the same order as the gravity anomaly. Hence it is permissible to disregard this term or some variation in the coefficient b_4 in the present problem, since doing so involves a discrepancy of the same order as the discrepancies involved in disregarding anomalies.

With this in mind, an attempt will be made to seek the field of force, the potential $\overline{V}$ of which would afford integration of the equations of motion of the satellite whilst being at the same time sufficiently near to the potential V.

We will introduce an inertial system of dimensionless curvilinear co-ordinates q_1, q_2, q_3.

The co-ordinates q_1, q_2 are determined from the formulae:

$$q_1 = \frac{2}{1}\left[\left(\frac{r^2}{d^2} - 1\right) + \sqrt{\left(\frac{r^2}{d^2} - 1\right)^2 + 4\frac{r^2}{d^2}\sin^2\phi}\right],$$

$$q_2 = \frac{1}{2}\left[\left(\frac{r^2}{d^2} - 1\right) - \sqrt{\left(\frac{r^2}{d^2} - 1\right)^2 + 4\frac{r^2}{d^2}\sin^2\phi}\right], \tag{2}$$

where d is some as yet arbitrary linear parameter.

The co-ordinate surfaces $q_1 = \text{const.}$ and $q_2 = \text{const.}$ are readily seen respectively as ellipsoids of rotation and single-sheet hyperboloids of rotation. The co-ordinate q_1, q_2 can vary over the range:

$$-1 \leqslant q_2 \leqslant 0 \leqslant q_1 \leqslant +\infty.$$

The co-ordinate q_3 represents the angle between the meridianal plane containing the given point and some initial plane passing through the earth's axis of rotation (longitude).

The system q_1, q_2, q_3 is not a system of elliptical co-ordinates in space in its usual definition[4]. We will, however, allowing some conditional approach, term the co-ordinates q_1, q_2 and q_3 dimensionless elliptical co-ordinates.

In K. Yakobi's classic work *Lectures on Dynamics*[5] it is observed that the equations of motion of a point in a field of force, the potential Π of which takes the form

$$\Pi = \frac{F(q_1) + \Phi(q_2)}{q_1 - q_2}, \tag{3}$$

where $F(q_1), \Phi(q_2)$ are arbitrary functions of the elliptical co-ordinates q_1, q_2, can always be integrated. We can state that the field with potential $\overline{V}$ equal to

$$\overline{V}=\frac{b_0}{d}\cdot\frac{\sqrt{q_1}}{q_1-q_2} \tag{4}$$

and consequently satisfying the condition (3) governing the selection of the parameter d, can be constructed throughout the entire outer space sufficiently near to the gravitational field of the total earth ellipsoid.

Assuming that the relationship d/r is small for all the considered values of r, the potential $\overline{V}$ is written in the form of the series:

$$\overline{V}=\frac{b_0}{r}\left[1-\frac{d^2}{r^2}P_2(\sin\phi)+\frac{d^4}{r^4}P_4(\sin\phi)-\ldots\right]. \tag{5}$$

The expansion (5) follows from equations (2) and (4). Taking the parameter d equal to

$$d=\sqrt{b_2/b_0} \tag{6}$$

the second term of the expansion (5) can be made equal to the second term of the expansion (1). (The first terms of these expansions coincide independently of the value d.)

The value of the linear parameter d thus selected is approximately 211 km ($b_0\approx 3{\cdot}99\times 10^{14}\frac{m^3}{\sec^2}$ $b_2\approx 1{\cdot}78\times 10^{25}\frac{m^5}{\sec^2}$). Consequently the ratio d/r in outer space will satisfy the inequality $\frac{d}{r}<\frac{1}{30}$ ($r>6370$ km), which justifies the assumption made above.

In the expansion (5), the terms containing the relationship d/r to the sixth and higher powers will constitute forces not exceeding fractions of a milligal, and can be disregarded. Hence the difference $\Delta V=\overline{V}-V$ can be described by the equation

$$\Delta V=\frac{b_0d^4-b_4}{r^5}P_4(\sin\phi)=\left(\frac{b_2^2}{b_0}-b_4\right)\cdot\frac{1}{r^5}P_4(\sin\phi).$$

For values of the radial (Δg_r) and meridianal (Δg_m) components of the acceleration, set up by the potential ΔV, the evaluation will apply $\left(b_4\approx 1{\cdot}6\times 10^{36}\frac{m^7}{\sec^2}\right)$:

$$|\Delta g_r|\leqslant 6{\cdot}1\text{ mgl}\quad|\Delta g_m|\leqslant 3{\cdot}3\text{ mgl}$$

Hence the potential $\overline{V}$ with the parameter d chosen in this way differs negligibly from the potential V. The maximum discrepancy in determining the gravitational force of the total earth ellipsoid using formula (4) does not exceed 7 mgl at the earth's surface. This discrepancy diminishes with removal from the earth.

It should be emphasised that the values Δg_r and Δg_m can be considered as discrepancies only if formula (4) is employed for describing the field of the total earth ellipsoid. If however the actual field of the earth's gravity is considered and the field with potential $\overline{V}$ is selected on the basis of its fresh normal section, then the potential ΔV must be simply included with the corresponding sign in the anomaly field potential. Further discrepancies should not naturally be incurred by this redistribution.

It should incidentally be observed that one further field of sufficiently simple structure may be indicated, closely approximating to the field of the total earth ellipsoid. The potential of this field V' is described by the formula:

$$V'=b_0/d\ .\ \text{arctg}\ 1/\sqrt{q_1} \tag{7}$$

The parameter d should now be chosen from the condition:

$$d=\sqrt{3b_2/b_0}.$$

The equivalent surfaces of the field of potential V' form confocal ellipsoids of rotation. Expansion of the potential V' in Legendre polynomials takes the form:

$$V'=\frac{b_0}{r}\left[1-\frac{1}{3}\frac{d^2}{r^2}P_2(\sin\phi)+\frac{1}{5}\frac{d^4}{r^4}P_4(\sin\phi)-\ldots\right].$$

The gravitational field of forces of the total earth ellipsoid will be described by formula (7) with a discrepancy not exceeding 1·5 mgl in outer space. This result can probably have some interest in itself.

The equations of motion of the satellite mass centre will now be written in canonical form. We will study the satellite path relative to an inertial system of co-ordinates. For the canonical variables we select the co-ordinates q_1, q_2,

q_3 and the corresponding generalized impulses

$$p_1=\frac{\partial T}{\partial \dot q_1};\ p_2=\frac{\partial T}{\partial \dot q_2};\ p_3=\frac{\partial T}{\partial \dot q_3}\quad \text{where } T=$$

$T(q_1, q_2, q_3, \dot q_1, \dot q_2, \dot q_3)$ is the kinetic energy of the satellite (the dot signifies time differentiation).

The canonical equations of motion in the Hamilton form take the form[4, 6]

$$\left.\begin{aligned}
\dot q_1&=\frac{\partial}{\partial p_1}(\hat T-\overline V),\quad \dot p_1=-\frac{\partial}{\partial q_1}(\hat T-\overline V),\\
\dot q_2&=\frac{\partial}{\partial p_2}(\hat T-\overline V),\quad \dot p_2=-\frac{\partial}{\partial q_2}(\hat T-\overline V),\\
\dot q_3&=\frac{\partial}{\partial p_3}(\hat T-\overline V),\quad \dot p_3=-\frac{\partial}{\partial q_3}(\hat T-\overline V),
\end{aligned}\right\}\qquad (8)$$

The function $\hat T=\hat T(q_1, q_2, q_3, p_1, p_2, p_3)$ represents the kinetic energy of the satellite expressed as canonical variables.

We will resolve the right-hand parts of equation (8). To this end we seek the relationship of the kinetic energy T to the co-ordinates q_1, q_2, q_3 and the generalized velocities $\dot q_1$, $\dot q_2$, $\dot q_3$. In the co-ordinates r, ϕ, q_3 the value of T is expressed by the formula:

$$T=\frac{1}{2}\left[\dot r^2+r^2\dot\phi^2+r^2\cos^2\phi\,\dot q_3^2\right]$$

(the satellite mass is taken as unity).

Differentiating the equations:

$$\left.\begin{aligned}
r^2&=d^2(1+q_1+q_2),\\
\operatorname{tg}^2\phi&=-\frac{q_1q_2}{(1+q_1)(1+q_2)},
\end{aligned}\right\}\qquad (9)$$

following from the relationships (2), we obtain:

$$\left.\begin{aligned}
\dot r^2&=\frac{d^2(\dot q_1+\dot q_2)^2}{4(1+q_1+q_2)},\\
r^2\dot\phi^2&=\frac{d^2[q_2(1+q_2)\dot q_1+q_1(1+q_1)\dot q_2]^2}{4q_1q_2(1+q_1)(1+q_2)(1+q_1+q_2)},\\
r^2\cos^2\phi\,\dot q_3^2&=d^2(1+q_1)(1+q_2)\dot q_3^2.
\end{aligned}\right\}\qquad (10)$$

After elementary transformations we arrive at the formula:

$$T=\frac{d^2}{8}\left[\frac{q_1-q_2}{q_1(1+q_1)}\dot q_1^2+\frac{q_2-q_1}{q_2(1+q_2)}\dot q_2^2+4(1+q_1)(1+q_2)\dot q_3^2\right].\qquad (11)$$

Thence follow the following expressions for the generalized impulses:

$$\left.\begin{aligned}
p_1&=\frac{d^2(q_1-q_2)}{4q_1(1+q_1)}\dot q_1,\\
p_2&=\frac{d^2(q_2-q_1)}{4q_2(1+q_2)}\dot q_2,\\
p_3&=d^2(1+q_1)(1+q_2)\dot q_3.
\end{aligned}\right\}\qquad (12)$$

Expressing the generalized velocities by impulses, and substituting in (11), we obtain:

$$\hat T=\frac{2}{d^2}\left[\frac{q_1(1+q_1)}{q_1-q_2}p_1^2+\frac{q_2(1+q_2)}{q_2-q_1}p_2^2+\frac{1}{4(1+q_1)(1+q_2)}p_3^2\right].\qquad (13)$$

As a result, the canonical equations of motion of the satellite in the inertial system of co-ordinates q_1, q_2, q_3 take the form:

$$\left.\begin{aligned}
\dot q_1&=\frac{4q_1(1+q_1)}{d^2(q_1-q_2)}p_1,\\
\dot q_2&=\frac{4q_2(1+q_2)}{d^2(q_2-q_1)}p_2,\\
\dot q_3&=\frac{1}{d^2(1+q_1)(1+q_2)}p_3,\\
\dot p_1&=-\frac{2}{d^2}\left[p_1^2-\frac{q_2(1+q_2)}{(q_1-q_2)^2}(p_1^2-p_2^2)-\frac{1}{4(1+q_1)^2(1+q_2)}p_3^2\right]-\frac{b_0}{d}\cdot\frac{q_1+q_2}{2\sqrt{q_1}(q_1-q_2)^2},\\
\dot p_2&=-\frac{2}{d^2}\left[p_2^2-\frac{q_1(1+q_1)}{(q_2-q_1)^2}(p_2^2+p_1^2)-\frac{1}{4(1+q_1)(1+q_2)^2}p_3^2\right]+\frac{b_0}{d}\cdot\frac{\sqrt{q_1}}{(q_1-q_2)^2},\\
\dot p_3&=0.
\end{aligned}\right\}\qquad (14)$$

Integrating the systems (14) it is possible, employing the Yakobi method[4-6], to substitute the search of a full integral by the following differential equation in partial derivatives:

$$\frac{2}{d^2}\left[\frac{q_1(1+q_1)}{q_1-q_2}\left(\frac{dW}{\partial q_1}\right)^2+\frac{q_2(1+q_2)}{q_2-q_1}\left(\frac{dW}{\partial q_2}\right)^2+\right.$$
$$\left.+\frac{1}{4(1+q_1)(1+q_2)}\left(\frac{\partial W}{\partial q_3}\right)^2\right]-$$
$$-\frac{b_0}{d}\cdot\frac{\sqrt{q_1}}{q_1-q_2}=C_1. \qquad (15)$$

Here $W=W(q_1, q_2, q_3)$ is the function sought, carrying the term "characteristic", $C_1=$const. is the total satellite energy.

Equation (15) is obtained, as is known, by substitution, in the left-hand side of the energy integral $T-\overline{V}=C_1$, the impulses p_1, p_2, p_3 respectively by $\frac{\partial W}{\partial q_1}, \frac{\partial W}{\partial q_2}, \frac{\partial W}{\partial q_3}$.

The complete integral of equation (15) will be the function $W(q_1, q_2, q_3, C_1, C_2, C_3)$, satisfying this equation and containing, in addition to the additive, two arbitrary constants C_2, C_3.

Equation (15) is re-written in the form:

$$q_1(1+q_1)\left(\frac{\partial W}{\partial q_1}\right)^2-q_2(1+q_2)\left(\frac{\partial W}{\partial q_2}\right)^2+$$
$$+\frac{q_1-q_2}{4(1+q_1)(1+q_2)}\left(\frac{\partial W}{\partial q_3}\right)^2-$$
$$\frac{b_0d}{2}\sqrt{q_1}=\frac{C_1d^2}{2}\ (q_1=q_2). \qquad (16)$$

It follows from the form of equation (16) that the function W can be presented as follows:

$$W=W_1(q_1, q_2, C_1, C_3)+C_2q_3. \qquad (17)$$

We then obtain the following equation, which the function W_1 must satisfy:

$$q_1(1+q_1)\left(\frac{\partial W_1}{\partial q_1}\right)^2-q_2(1+q_2)\left(\frac{\partial W_1}{\partial q_2}\right)^2+$$
$$+\frac{C_2^2(q_1-q_2)}{4(1+q_1)(1+q_2)}-\frac{b_0d}{2}\sqrt{q_1}=\frac{C_1d^2}{2}(q_1-q_2). \qquad (18)$$

Taking into account the obvious equation

$$\frac{q_1-q_2}{(1+q_1)(1+q_2)}=\frac{1}{1+q_2}-\frac{1}{1+q_1},$$

we can re-write equation (18) in the form:

$$q_1(1+q_1)\left(\frac{\partial W_1}{\partial q_1}\right)^2-\frac{C_2^2}{4(1+q_1)}-\frac{b_0d}{2}\sqrt{q_1}-$$
$$\frac{C_1d^2}{2}q_1=q_2(1+q_2)\left(\frac{\partial W_1}{\partial q_2}\right)^2-$$
$$\frac{C_2^2}{4(1+q_2)}-\frac{C_1d^2}{2}q_2. \qquad (19)$$

This latter equation is broken down into two:

$$q_1(1+q_1)\left(\frac{\partial W_1}{\partial q_1}\right)^2-\frac{C_2^2}{4(1+q_1)}-$$
$$\frac{b_0d}{2}\sqrt{q_1}-\frac{C_1d^2}{2}q_1=C_3,$$
$$q_2(1+q_2)\left(\frac{\partial W_1}{\partial q_2}\right)^2-\frac{C_2^2}{4(1+q_2)}-\frac{C_1d^2}{2}q_2=C_3. \qquad (20)$$

It can readily be seen that the function W_1, satisfying the relationships (20), will be a full integral of equation (19).

Resolving equations (20) with reference to $\frac{\partial W_1}{\partial q_1}$ and $\frac{qW_1}{\partial q_2}$

$$\frac{\partial W_1}{\partial q_1}=\sqrt{\frac{\frac{C_2^2}{4(1+q_1)}+\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3}{q_1(1+q_1)}}$$
$$\frac{\partial W_1}{\partial q_2}=\sqrt{\frac{\frac{C_2^2}{4(1+q_2)}+\frac{C_1d^2}{2}q_2+C_3}{q_2(1+q_2)}} \qquad (21)$$

and integrating, we obtain, taking into account the relationship (17), the following expression for the characteristic function:

$$W=\int dq_1\sqrt{\frac{\frac{C_2^2}{4(1+q_1)}+\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3}{q_1(1+q_1)}}+$$
$$+\int dq_3\sqrt{\frac{\frac{C_2^2}{4(1+q_2)}+\frac{C_1d^2}{2}q_2+C_3}{q_2(1+q_2)}}+C_2q_3. \qquad (22)$$

After finding the function W, the integrals of the canonical equations of motion (14) may be given by the equations:

$$\frac{\partial W}{\partial C_2}=C_4,\quad \frac{\partial W}{\partial C_3}=C_5,$$
$$\frac{\partial W}{\partial q_1}=p_1,\quad \frac{\partial W}{\partial q_2}=p_2,\quad \frac{\partial W}{\partial q_3}=p_3, \tag{23}$$
$$\frac{\partial W}{\partial C_1}=t+C_6,$$

where C_4, C_5, C_6 are arbitrary constants and t the time.

The first two integrals of system (23) are geometrical; they determine the shape of the satellite trajectory. The last integral is the kinematic: it applies the law of motion of the satellite along the trajectory. The remaining integrals serve for determining the impulses (velocities). In the developed form the system of integrals (23) will appear as follows:

$$\left.\begin{aligned}&\int\frac{C_2dq_1}{4(1+q_1)\sqrt{q_1\left[\frac{C_2^2}{4}+(1+q_1)\left(\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3\right)\right]}}+\\&+\int\frac{C_2dq_2}{4(1+q_2)\sqrt{q_2\left[\frac{C_2^2}{4}+(1+q_2)\frac{C_1d^2}{2}q_2+C_3\right)\right]}}+q_3=C_4,\end{aligned}\right| \tag{24}$$

$$\left.\begin{aligned}&\int\frac{dq_1}{2\sqrt{q_1\left[\frac{C_2^2}{4}+(1+q_1)\left(\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3\right)\right]}}+\\&+\int\frac{dq_2}{2\sqrt{q_2\left[\frac{C_2^2}{4}+(1+q_2)\left(\frac{C_1d^2}{2}q_2+C_3\right)\right]}}=C_5,\end{aligned}\right| \tag{25}$$

$$\left.\begin{aligned}p_1&=\sqrt{\frac{\frac{C_2^2}{4(1+q_1)}+\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3}{q_1(1+q_1)}},\end{aligned}\right| \tag{26}$$

$$\left.\begin{aligned}p_2&=\sqrt{\frac{\frac{C_2^2}{4(1+q_2)}+\frac{C_1d^2}{2}q_2+C_3}{q_2(1+q_2)}},\end{aligned}\right| \tag{27}$$

$$p_3=C_2, \tag{28}$$

$$\left.\begin{aligned}&\int\frac{q_1\cdot d^2\cdot dq_1}{4\sqrt{q_1\left[\frac{C_2^2}{4}+(1+q_1)\left(\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_2+C_3\right)\right]}}+\\&+\int\frac{q_2\cdot d^2\cdot dq_2}{4\sqrt{q_2\left[\frac{C_2^2}{4}+(1+q_2)\left(\frac{C_1d^2}{2}q_1+C_3\right)\right]}}=t+C_6.\end{aligned}\right| \tag{29}$$

Equations (24)–(29) completely solve the problem of the satellite path under the influence of a gravitational field of force with potential $\overline{V}$.

We will reduce the results obtained to a form more suitable for practical application. Assume a satellite path at the initial moment*:

$$t=t_k=0,$$
$$q_1=q_{1k},\ q_2=q_{2k},\ q_3=q_{3k},$$
$$p_1=p_{1k},\ p_2=p_{2k},\ p_3=p_{3k}.$$

Inserting the initial values of the variables in (24), (25), (29), we can write:

$$q_3=q_{3k}-\frac{1}{4}C_2\left\{\int_{q_{1k}}^{q_1}\frac{dq_1}{(1+q_1)\sqrt{q_1\left[\frac{C_2^2}{4}+(1+q_1)\left(\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3\right)\right]}}+\right.$$
$$\left.+\int_{q_{2k}}^{q_2}\frac{dq_2}{(1+q_2)\sqrt{q_2\frac{C_2^2}{4}+(1+q_2)\left(\frac{C_1d^2}{2}q_2+C_3\right)}}\right\}, \tag{30}$$

$$\left.\int_{q_{1k}}^{q_1}\frac{dq_1}{\sqrt{q_1\left[\frac{C_2^2}{4}+(1+q_1)\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3\right)}}+\right.$$
$$\left.+\int_{q_{2k}}^{q_2}\frac{dq_2}{\sqrt{q_2\left[\frac{C_2^2}{4}+(1+q_2)\frac{C_1d^2}{2}q_2+C_3\right)}}=0,\right. \tag{31}$$

$$t=\frac{d^2}{4}\left\{\int_{q_{1k}}^{q_1}\frac{q_1dq_1}{\sqrt{q_1\frac{C_2^2}{4}+(1+q_1)\frac{b_0d}{2}\sqrt{q_1}+\frac{C_1d^2}{2}q_1+C_3\right)}}+\right.$$
$$\left.+\int_{q_{2k}}^{q_2}\frac{q_2dq_2}{\sqrt{q_2\left[\frac{C_2^2}{4}+(1+q_2)\frac{C_1d^2}{2}q_2+C_3\right)}}\right\}. \tag{32}$$

To find the arbitrary constants C_1, C_2, C_3, the constant C_1 (total satellite energy) is expressed by the initial conditions as follows:

$$C_1=\hat{T}_k-\overline{V}_k=\frac{2}{d^2}\left[\frac{q_{1k}(1+q_{1k})}{q_{1k}-q_{2k}}p_{1k}^2+\right.$$
$$\left.+\frac{q_{2k}(1+q_{2k})}{q_{2k}-q_{1k}}p_{2k}^2+\frac{1}{4(1+q_{1k})(1+q_{2k})}p_{3k}^2\right]-$$
$$-\frac{b_0}{d}\cdot\frac{\sqrt{q_{1k}}}{q_{1k}-q_{2k}}. \tag{33}$$

The constant C_2 obviously equals:

$$C_2=p_{3k}. \tag{34}$$

In order to determine the constant C_3 use is made of equation (27).

Considering this equation at the initial point and resolving it with reference to C_3 we arrive at the relationship:

* The index k in future corresponds to the initial point throughout.

$$C_3=\frac{b_0 d}{2}\cdot\frac{q_{2k}\sqrt{q_{1k}}}{q_{1k}-q_{2k}}-\frac{p_{2k}^2(1+q_{1k}+q_{2k})}{4(1+q_{1k})(1+q_{2k})}-$$

$$\frac{q_{1k}q_{2k}}{q_{1k}-q_{2k}}[p_{1k}^2(1+q_{1k})-p_{2k}^2(1+q_{2k})]. \qquad (35)$$

We pass now to the dimensionless arbitrary constants D_1, D_2, D_3, governed by the relationships:

$$D_1=C_1\cdot\frac{d}{b_0};\ D_2=\frac{C_2}{\sqrt{b_0 d}};\ D_3=\frac{C_3}{b_0 d}. \qquad (36)$$

Simultaneous substitution is applied for the variables q_1, q_2 according to the formulae:

$$\xi=\sqrt{q_1},\qquad \eta=\sqrt{-q_2}. \qquad (37)$$

The integrals of the equations of motion of the satellite can then be written in the form:

$$q_3=q_{3k}+\frac{D_2}{\sqrt{2}}[I_2(\eta)-I_1(\xi)], \qquad (38)$$

$$I_3(\xi)=I_4(\eta), \qquad (39)$$

$$t=\sqrt{\frac{d^3}{2b_0}}[I_5(\xi)+I_6(\eta)], \qquad (40)$$

$$p_1=\sqrt{\frac{b_0 d}{2}}\cdot\frac{\sqrt{Q(\xi)}}{\xi(1+\xi^2)}, \qquad (41)$$

$$p_2=-\sqrt{\frac{b_0 d}{2}}\cdot\frac{\sqrt{P(\eta)}}{\eta(1-\eta^2)}, \qquad (42)$$

$$p_3=D_3\sqrt{b_0 d}. \qquad (43)$$

where:

$$I_1(\xi)=\int_{\xi_k}^{\xi}\frac{d\xi}{(1-\xi^2)\sqrt{Q(\xi)}},$$

$$I_2(\eta)=\int_{\eta_k}^{\eta}\frac{d\eta}{(1-\eta^2)\sqrt{P(\eta)}},$$

$$I_3(\xi)=\int_{\xi_k}^{\xi}\frac{d\xi}{\sqrt{Q(\xi)}},$$

$$I_4(\eta)=\int_{\eta_k}^{\eta}\frac{d\eta}{\sqrt{P(\eta)}},$$

$$I_5(\xi)=\int_{\xi_k}^{\xi}\frac{\xi^2 d\xi}{\sqrt{Q(\xi)}},$$

$$I_6(\eta)=\int_{\eta_k}^{\eta}\frac{\eta^2 d\eta}{\sqrt{P(\eta)}}.$$

The polynomials of the fourth order $Q(\xi)$ and $P(\eta)$ are determined by the formulae:

$$Q(\xi)=m\xi^4+\xi^3+n\xi^2+\xi+s,$$

$$P(\eta)=-m\eta^4+n\eta^2-s.$$

The coefficients of the polynomials m, n, s are the following functions of the arbitrary constants D_1, D_2, D_3:

$$m=D_1,\ n=D_1+2D_3,\ s=\frac{1}{2}D_2^2+2D_3.$$

The integrals I_1, I_2, I_3, I_4, I_5, I_6 are elliptical integrals. Without going into detail as to the method of calculation, it will be said that in carrying out the calculation on a computer it is more convenient to employ expansions of these integrals into series according to powers of small parameters.

Comparative calculations have shown that the programme employing the formulae obtained affords a reduction, in the ratio of tens and hundreds (as a function of the extent of the calculated section of the orbit and the number of points involved in the determination), in the computer time necessary for carrying out the calculations as compared with the programme based on mathematical integration of the equations of motion.

It is interesting to dwell on some particular cases of the considered motion, yielding more simply to qualitative analysis.

In the first place we will write the formulae relating the generalized impulses p_1, p_2, p_3 to the satellite velocity projections v_1, v_2, v_3 along the positive co-ordinate lines of the spherical inertial system of co-ordinates r, ϕ, q_3. Employ-

ing equations (10) and (12) and considering that

$$v_1=\dot{r},\quad v_2=r\dot{\phi},\quad v_3=r\cos\phi\dot{q}_3,$$

we obtain:

$$v_1=\frac{2}{r(q_1-q_2)}[q_1(1+q_1)p_1-q_2(1+q_2)p_2],$$

$$v_2=\frac{r\sin 2\phi}{d^2(q_2-q_1)}(p_2-p_1),$$

$$v_3=\frac{p_3}{d\sqrt{(1+q_1)(1+q_2)}}. \tag{44}$$

Expressing explicitly the impulses p_1, p_2, p_3 by the velocities v_1, v_2, v_3 we have:

$$p_1=\frac{r}{4q_1(1+q_1)}\left[2v_1\frac{q_1(1+q_1)}{1+q_1+q_2}+v_2\sin 2\phi\right],$$

$$p_2=\frac{r}{4q_2(1+q_2)}\left[2v_1\frac{q_2(1+q_2)}{1+q_1+q_2}+v_2\sin 2\phi\right],$$

$$p_3=d\sqrt{(1+q_1)(1+q_2)}\cdot v_3. \tag{45}$$

We will express also the arbitrary constants D_1, D_2, D_3 by the quantities v_{1k}, v_{2k}, v_{3k}. Considering the equations (45) at the initial point and inserting the expression for p_{1k}, p_{2k}, p_{3k} in formulae (33), (34), (35), we obtain:

$$D_1=\frac{d\cdot v_k^2}{2b_0}-\frac{\sqrt{q_1^k}}{q_{1k}-q_{2k}},$$

$$D_2=\frac{v_{3k}r^*\cos\phi_k}{\sqrt{b_0d}}=v_{3k}\sqrt{\frac{d}{b_0}}\cdot\sqrt{(1+q_{1k})(1+q_{2k})},$$

$$D_3=\frac{d}{4b_0}\left[\frac{2b_0}{d}\cdot\frac{q_{2k}\sqrt{q_{1k}}}{q_{1k}-q_{2k}}-(1+q_{1k}+q_{2k})(v_{2k}^2+v_{3k}^2)+(v_{1k}\sin\phi_k+v_{2k}\cos\phi_k)^2\right], \tag{46}$$

where

$$v_k^2=v_{1k}^2+v_{2k}^2+v_{3k}^2.$$

We will first analyze the case $D_1=D_2=D_3=0$. The equation $D_1=0$ signifies that the total "satellite" energy possesses that minimum quantity which is necessary for removing the "satellite" from the earth to infinity*. It follows from the relationship $D_2=0$ that the motion occurs in the meridianal plane. The equation $D_3=0$ for $D_1=0$ and $v_{3k}=0$ involves the following condition:

$$\frac{v_{2k}}{v_{1k}}=\frac{1+q_{2k}}{q_{1k}}tg\,\phi_k. \tag{47}$$

It can readily be seen that on satisfying the condition (47) the "satellite" initial velocity vector will be directed normal to the ellipsoid of rotation, governed by the equation $q_1=q_{1k}$. The polynomial $Q(\xi)$ takes the form:

$$Q(\xi)=\xi(1+\xi^2).$$

For the impulse p_1 the following expression will apply:

$$p_1=\sqrt{\frac{b_0d}{2}}\cdot\frac{1}{\sqrt{\xi(1+\xi^2)}}. \tag{48}$$

The polynomial $P(\eta)$ is identically equal to zero. For $q_{2k}\neq 0$ and $q_{2k}\neq -1$, from the integral (42) it follows directly that:

$$p_2=0$$

and consequently

$$q_2=q_{2k}=\text{const}.$$

With $q_{2k}=0$ (the initial point lies in the equatorial plane) or $q_{2k}=-1$ (the initial point lies on the earth's axis of rotation), when in the right-hand side of formula (42) there is an indeterminate, this result involves a study of the derivative

$$\frac{dq_2}{dq_1}=\frac{q_2(1+q_2)}{q_1(1+q_1)}\cdot\frac{p_2}{p_1}=-\frac{\eta\sqrt{P(\eta)}}{\xi\sqrt{Q(\xi)}}.$$

Consequently, with the departure of the "satellite" from the earth to infinity at the minimum initial velocity necessary for this in a direction normal to the ellipsoid $q_1=q_{1k}$, the path will follow a hyperbola (focal length $2d$), the vertical axis of which coincides with the earth's axis of rotation, whilst the real axis lies in the meridian plane of the initial point. This case is analogous to the parabolic case in the theory of Kepler motion.

* The term " satellite " takes an arbitrary character in this case.

The velocity components v_1, v_2 and time of flight t with $D_1=D_2=D_3=0$ are associated with the ξ co-ordinate by the relationships:

$$v_1=\sqrt{\frac{2b_0}{d}}\cdot\frac{\xi}{\xi^2+\eta_k^2}\sqrt{\frac{\xi(1+\xi^2)}{1+\xi^2-\eta_k^2}},$$

$$v_2=\sqrt{\frac{2b_0}{d}}\cdot\frac{\eta_k}{\xi^2+\eta_k^2}\sqrt{\frac{\xi(1-\eta_k^2)}{1+\xi^2-\eta_k^2}},$$

$$t=\sqrt{\frac{d^3}{2b_0}}\int_{\xi_k}^{\xi}\frac{(\xi^2+\eta_k^2)\,d\xi}{\sqrt{\xi(1+\xi^2)}}.$$

These formulae follow directly from (44), (48) and the expression for the generalized velocity $\dot{q}_1$ (see (14)):

$$\dot{q}_1=2\sqrt{\frac{2b_0}{d^3}}\cdot\frac{\xi\sqrt{\xi(1+\xi^2)}}{\xi^2+\eta_k^2}.$$

We will pass now to the analysis of the equations of motion for the case where the constants D_1, D_2, D_3 are connected by the relationships:

$$D_2^2=2(1+\xi_k^2)^2\left(D_1+\frac{1}{2\xi_k}\right),$$
$$D_3=-\frac{1}{2}D_1(1+2\xi_k^2)-\frac{1+3\xi_k^2}{4\xi_k}. \tag{49}$$

It can readily be confirmed that, with such limitations imposed on the initial conditions, the satellite path will proceed along the surface of an ellipsoid $q_1=q_{1k}$. By analogy with the circular satellite we will call this satellite ellipsoidal.

In actual fact, inserting in the polynomial $Q(\xi)$ and in its first and second derivatives $Q'(\xi)$ and $Q''(\xi)$ the value $\xi=\xi_k$ and expressing n, s from (49) by D_1 and ξ_k, we obtain:

$$Q(\xi_k)=0,\qquad Q'(\xi_k)=0,$$
$$Q''(\xi_k=8D_1\xi_k^2+\frac{3\xi_k^2-1}{\xi_k}.$$

The polynomial $P(\eta)$, with $\eta=\eta_k$ and the same values m, n, s, is equal to:

$$P(\eta_k)=-D_1(\xi_k^2+\eta_k^2)^2-\frac{(\xi_k^2+\eta_k^2)^2+(1-\eta_k^2)(\eta_k^2-\xi_k^2)}{2\xi_k}$$

Since the inequalities must always apply

$$P(\eta_k)\geqslant 0;\qquad D_2^2\geqslant 0,$$

the constant D_1 is contained within the range

$$-\frac{1}{2\xi_k}\leqslant D_1\leqslant-\frac{1}{2\xi_k}\left[1-\frac{(1-\eta_k^2)(\xi_k^2-\eta_k^2)}{(\xi_k^2+\eta_k^2)^2}\right]. \tag{50}$$

The quantity

$$\Delta=\frac{(1-\eta_k^2)(\xi_k^2-\eta_k^2)}{(\xi_k^2+\eta_k^2)^2}$$

cannot fall below zero, since

$$\xi_k^2\cong\frac{r_k^2}{d^2}\geqslant 1\geqslant\eta_k^2.$$

(It is assumed that the initial point occurs above the earth's surface). Hence the inequalities (50) will always be consistent. Employing (50), we arrive at the relationship

$$Q''(\xi_k)\leqslant-\frac{1-\xi_k^2}{\xi_k}+\frac{4\xi_k(1-\eta_k^2)(\xi_k^2-\eta_k^2)}{(\xi_k^2+\eta_k^2)^2}\leqslant$$
$$\leqslant-\frac{1+\xi_k^2}{\xi_k(\xi_k^2+\eta_k^2)^2}[(\xi_k^2+\eta_k^2)^2-4\xi_k^2(1-\eta_k^2)]<0,$$

since $\xi_k\geqslant 2$. This means that in the vicinity of point ξ_k with $\xi\geq\xi_k$, the polynomial $Q(\xi)$ takes on negative values. Hence the motion is possible only on condition that the co-ordinate ξ is identically equal to ξ_k, that is to say along the surface of the ellipsoid $q_1=q_{1k}$. With $D_1=(1/2\xi_k)$ the path must proceed in the meridian plane, since on the basis of (49), $D_2=0$. In this case the satellite trajectory forms an ellipse (focal length $2d$) the short axis of which coincides with the earth's axis of rotation, whilst the long axis lies on the meridian plane, passing through the initial point. The long semi-axis of this ellipse a and the eccentricity e are expressed by the co-ordinate ξ_k and the parameter d in the following manner:

$$a=d\sqrt{1+\xi_k^2},$$
$$e=\frac{1}{\sqrt{1+\xi_k^2}}.$$

The time of flight of this satellite is connected with the co-ordinate η by the relationship

$$t=\sqrt{\frac{d^2\xi_k}{b_0}}\int_{\eta^k}^{\eta}\frac{(\xi_k^2+\eta^2)\,d\eta}{\sqrt{(1-\eta^2)(\xi_k^2-\eta^2)}}\,.$$

With $D_1=-(1/2\xi_k)(1-\Delta)$ and $q_{2k}=0$ we obtain a circular satellite with an orbit lying in the equatorial plane.

The height of the ellipsoidal satellite above the surface of the common earth ellipsoid will increase with displacement of the satellite from the equator towards the pole. With $D_2=0$ the maximum difference in heights for an orbit at a fairly small distance from the earth is approximately 18 km.

In the general case, it is possible, for each given position of the initial point, to establish an unambiguous relationship between the azimuth of the plane of projection A_k and the velocity v_k necessary for obtaining an ellipsoidal satellite. The inverse relationship will not be unambiguous, since the same velocity will obviously correspond to directions with azimuths $\pm A_k$ and $(180^\circ\pm A_k)$. These relationships follow from the relationships (46) and (49). Given, for example, a value v_k and having determined the value D_1, it is possible to find D_2 from (49) and consequently also v_{3k}. The A_k azimuth is determined from the formula:

$$\sin A_k=\frac{v_{3k}}{v_k}\,.$$

The A_k azimuth here signifies the angle between the direction of the meridian and the initial velocity vector in a plane tangential to the ellipsoid $q_1=q_{1k}$. (For the ellipsoidal satellite the initial velocity vector always lies in this plane).

The longitude and time of flight of the ellipsoidal satellite, in the general case, is equal to:

$$q_3=q_{3k}-\frac{D_2}{\sqrt{2}}\left[\frac{1}{1+\xi_k^2}I_4(\eta)-I_2(\eta)\right]\,.$$

$$t=\sqrt{\frac{d^2}{2b_0}}\,[\xi_k^2 I_4(\eta)+I_6(\eta)].$$

The inequalities (50) and relationships (49) can be treated as conditions governing the minimum velocity $v_{k,\min}$, necessary for the existence of a satellite (understanding by the latter a freely projected body, the co-ordinate q_1 of which satisfies the inequality $q_1\geqslant q_{1k}$ throughout the whole time of travel). This velocity is a function not only of the distance to the earth's centre r_k but also of the latitude ϕ_k and the azimuth of the plane of projection A_k. For a given value q_{1k} the velocity $v_{k,\min}$ has the greatest value for $q_{2k}=0$ and $A_k=\pm 90^\circ$, i.e. existence of a circular satellite, travelling in the equatorial plane.

A few words should be said about the possibility of "satellite" motion in a field with potential $\overline{V}$ with constant q_2 co-ordinate. It was pointed out above that this case will occur in particular with $D_1=D_2=D_3=0$. We will state the problem in a more general manner: under what conditions must the constants D_1, D_2, D_3 be related in order to maintain the equality throughout the whole time of flight

$$q_2=q_{2k}=\text{const.}$$

Excluding from consideration for the present the cases $q_{2k}=0$ and $q_{2k}=-1$, in the energy integral

$$\frac{2}{b_0 d}\left[\frac{\xi^2(1+\xi^2)}{\xi^2+\eta^2}p_1^2+\frac{\eta^2(1-\eta^2)}{\xi^2+\eta^2}p_2^2+\right.$$
$$\left.+\frac{1}{4(1+\xi^2)(1-\eta^2)}p_3^2\right]-\frac{\xi}{\xi^2+\eta^2}=D_1$$

assume $q_2=q_{2k}$; $p_2=0$; $p_3=D_2\sqrt{b_0 d}$. Since with $q_2=\text{const.}$ we obtain $p_2=0$ then, considering the fifth equation of the system (14), we have

$$p_1^2\frac{\xi^2(1+\xi^2)}{\xi^2+\eta_k^2}=\frac{b_0 d}{2}\left[\frac{\xi}{\xi^2+\eta_k^2}+\right.$$
$$\left.+\frac{D_2^2(\xi^2+\eta_k^2)}{2(1+\xi^2)(1-\eta_k^2)^2}\right]\,.$$

Writing the expression for p_1 in the energy integral, we arrive at the following relationship between D_1 and D_2, not containing the running co-ordinate ξ:

$$D_1=\frac{D_2^2}{2(1-\eta_k^2)^2}\qquad(51)$$

Comparing (35), expressing D_3 in terms of D_2:

$$D_3 = -\frac{D_2^2(1-2\eta_k^2)}{4(1-\eta_k^2)^2} \quad . \tag{52}$$

It thus follows that the condition $q_2 = q_{2k} =$ const. cannot be realized with $D_1 < 0$ (in this case the total "satellite" energy is inadequate for departure from earth to infinity). With $D_1 = 0$ we have $D_2 = 0$, $D_3 = 0$, i.e. the case already examined. With $D_1 > 0$, motion with constant value of the co-ordinate q_2 will be attainable, since the initial conditions can be connected by the relationships (51), (52). In this case the "satellite" trajectory will lie on the surface of a single-sheet hyperboloid of rotation $q_2 = q_{2k}$.

If $q_{2k} = 0$ and $q_{2k} = -1$, then it is evident that for any D_1 it is possible to obtain $q_2 = q_{2k} =$ const. In the first case it is necessary, for this, to assume $v_{2k} = 0$ (motion in the equatorial plane), in the second case, to direct the initial velocity vector along the vertical (motion along the earth's axis of rotation). It should be observed that the transformations employed in the development of formulae (50) and (51) are unsuitable for establishing the connection between D_1, D_2, D_3 for $q_{2k} = 0$ and $q_{2k} = -1$.

The analogy with the case $q_1 = q_{1k} =$ const. emerges quite definitely from the analysis carried through. This explains the nature of the elliptical co-ordinates employed in the solution.

The theory of satellite motion in a field with potential $\overline{V}$, stated in broad terms in the present paper, can be regarded as a further approximation as compared with the theory of Kepler motion* in the description of the real flight chart of an earth satellite in empty space. The potential $\overline{V}$, on which this theory is based, approaches considerably closer to the real field of earth gravitational forces, than the point mass potential, leaving out of account, by the nature of things, only the gravity anomaly. The main difference between the earth's field and the point mass field—a difference due to the ellipsoidal shape of the earth—enters fully in the potential $\overline{V}$.

In addition to employing the relationships obtained for purely calculation purposes, they can be used for solving problems regarding secular and periodic perturbations of the elements of elliptical orbits of artificial satellites under the influence of the earth's oblateness. In subsequent publications the author hopes to set out the results of investigations in this direction.

* The integrals of Kepler motion follow, as a particular case, with $d \to 0$ from the integrals (38)-(43).

REFERENCES

1. I. D. ZHONGOLOVISCH, *Trud. Inst. Teor. Fiz.*, **3** (1952).
2. A. A. MIKHAILOV, *Course on gravimetry and theory of the earth's form*, Main Geodesy and Cartography Administration Publishing House, Moscow (1939). (In Russian).
3. N. I. IDEL'SON, *Potential theory*, United Scientific Press (1936). (In Russian).
4. G. K. SUSLOV, *Theoretical Mechanics*. G.T.T.I., Moscow (1946). (In Russian).
5. K. YAKOBI, *Lectures on Dynamics*, G.T.T.I., Moscow (1936). (In Russian).
6. N. N. BUKHHOL'TS, *Basic Theoretical Mechanics Course*. **2**, G.T.T.I. Moscow (1939). (In Russian).

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 97–106. Printed in Great Britain.

PROBLEMS OF GUIDANCE IN INTERPLANETARY SPACE

B. V. RAUSHENBAKH and E. N. TOKAR'

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* No. 5, p. 41 (1960)

The present paper studies some of the problems involved in the guidance of an interplanetary missile. Interplanetary missile signifies a device travelling outside the atmospheric limits. It may be an artificial Earth satellite, one stage of a space rocket or the like. The guidance system for such a missile must provide for varying the path of the mass centre of the missile and for performing the desired rotations of the missile around its mass centre, together with maintaining the axes associated with the missile in a given position. We will consider below only the second of these problems—rotations of the missile and maintenance of a given orientation.

Two methods have been proposed by K. E. Tsiolkovskii for rotation of the missile around its mass centre, namely rocket thrust, the line of action of which does not pass through the mass centre and the use of reactive flywheels. Missile guidance employing jet engines can be fairly readily analyzed. This cannot be said, however, regarding guidance-employing reactive flywheels.

To generalize we will assume that the number of stabilizing flywheels in the missile body is equal to n. Conversion to any given number of flywheels does not in any way complicate the construction of the equations of motion. The arrangement of the flywheels and the orientation of their axes in the missile body will be taken as arbitrary, but invariable in time; we will also assume that the flywheels are balanced statically and dynamically on their axes; in other words their axes of rotation coincide with the main central axes of inertia. This assumption is obviously justified for axially symmetrical flywheels.

1. EQUATIONS OF MOTION AROUND THE MASS CENTRE

In order to draw up the equations of motion we will bring in two systems of co-ordinates taking as the common reference origin the point O—the mass centre of the material system consisting of the missile body and the reactive flywheels. In Fig. 1, $OX_0Y_0Z_0$ is the reference

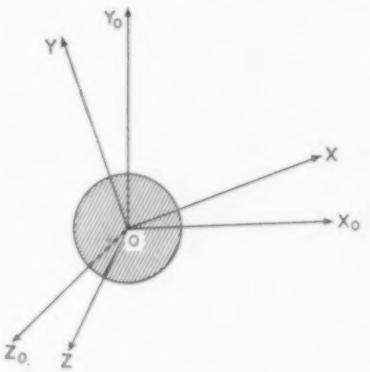


Fig. 1. Mutual position of the missile and reference systems $OXYZ$ and $OX_0Y_0Z_0$.

system, the axes of which remain constantly parallel to the axes of some inertial system, whilst $OXYZ$ is the reference system associated invariably with the missile body; the question of orientation of the OX, OY and OZ axes in the missile body will be meantime left open. The instantaneous angular velocity of the $OXYZ$ system relative to the $OX_0Y_0Z_0$ system will be termed ω; the projections of this angular velocity on the OX, OY and OZ axes will be taken respectively as p, q and r.

We will apply a positive direction to the supporting axes of each of the flywheels; the orientation of the flywheel axes will be determined by the cosines of the angles formed by the posi ive direction of the flywheel axes to the positive directions of the OX, OY and OZ axes. For the ith flywheel let these directional cosines be a_{ix}, a_{iy} and a_{iz}. The angular velocity of the ith flywheel relative to the $OXYZ$ system we will term $\boldsymbol{\omega}_i$; and the projection of this vector on the axis of rotation of the flywheel ω_i. The projections of the $\boldsymbol{\omega}_i$ vector on the OX,

OY and OZ axes with this system of notation will be written respectively $\omega_i a_{ix}$, $\omega_i a_{iy}$, $\omega_i a_{iz}$.

The kinetic moment of the given material system in relative motion around the mass centre, referred to point O, will be

$$\mathbf{K} = \int_S [\mathbf{rv}]\, dm$$

where $\mathbf{v}$ is the velocity of the mass element dm relative to the $OX_0Y_0Z_0$ system, and $\mathbf{r}$ is the radius vector of the given mass element dm; integration is carried out over the entire S system.

We will represent the $\mathbf{K}$ vector as the sum of the kinetic moment of the missile body and of the kinetic moments of the reactive flywheels, whilst the kinetic moments of the flywheels will be decomposed into the kinetic moments governed by the transfer velocities of the flywheel points in the $OX_0Y_0Z_0$ system, and into the kinetic moments governed by the velocities of self-rotation of the flywheels relative to the $OXYZ$ system:

$$\mathbf{K} = \int_C [\mathbf{rv}]\, dm + \sum_{i=1}^{n} \int_{M_i} [\mathbf{rv}_{tr}]\, dm + \sum_{i=1}^{n} \int_{M_i} [\mathbf{rv}_{rel}]\, dm$$

where, in the first integral, integration is carried out over the missile body C and in the remaining integrals over the ith flywheel M_i.

The velocity of any point in the body $\mathbf{v}$ (including the transfer velocity of each point of the flywheel $\mathbf{v}_{tr}$) is expressed by the angular velocity of the $OXYZ$ system according to the Euler formula $\mathbf{v} = [\omega \mathbf{r}]$. This affords grouping of the first $n+1$ summands.

$$\int_C [\mathbf{rv}]\, dm + \sum_{i=1}^{n} \int_{M_i} [\mathbf{rv}_{tr}]\, dm = \int_S [\mathbf{r}\,[\omega \mathbf{r}]]\, dm$$

hence

$$\mathbf{K} = \int_S [\mathbf{r}\,[\omega \mathbf{r}]]\, dm + \sum_{i=1}^{n} \int_{M_i} [\mathbf{rv}_{rel}]\, dm$$

We will study the second summand of the expression obtained.

Since the supporting axes of the flywheels in the missile body coincide with their main axes of inertia, the kinetic moments of relative motion of the flywheels in the $OXYZ$ system, taking the mass centre of the flywheels as the reduced centres, will be $\boldsymbol{\omega}_1 J_1$, $\boldsymbol{\omega}_2 J_2, \ldots \boldsymbol{\omega}_n J_n$. Here $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2 \ldots \boldsymbol{\omega}_n$ are the angular velocities of the flywheels introduced above, relative to the system $OXYZ$, whilst J_1, $J_2, \ldots J_n$ are their axial moments of inertia. In the motion of the flywheels relative to the $OXYZ$ system the velocities of the flywheel mass centres are zero (the flywheels are balanced on their support axes in the missile body), governed by the independence of the corresponding kinetic moments from the centre of reduction. Taking the point O as the common centre of reduction for all the flywheels we obtain

$$\int_{M_i} [\mathbf{rv}_{rel}]\, dm = \boldsymbol{\omega}_i J_i$$

whence the expression for $\mathbf{K}$ will be:

$$\mathbf{K} = \int_S [\mathbf{r}\,[\omega \mathbf{r}]]\, dm + \sum_{i=1}^{n} \boldsymbol{\omega}_i J_i \qquad (1)$$

Presenting the kinetic moment of the system in this form, the orientation of the OX, OY and OZ axes in the missile body can be selected in such a way as to afford the simplest expressions for the projection of the kinetic moment on the axes of the $OXYZ$ system. Determination of this system of axes is important not only from the standpoint of simplifying the construction of the equations of motion but also from the standpoint of improving the dynamics of the stabilized missile. Arranging the planes in which the signals of the sensitive elements of the stabilizing system are generated parallel to the co-ordinate planes of the $OXYZ$ system, suitably selected, and arranging the axes of the reactive flywheels parallel to the axes of the same system, we can reduce to a minimum the mutual influence of missile oscillations around the three axes of stabilization.

It can readily be seen that the first summand in the expression (1) represents the kinetic moment which the system would possess if the

flywheels were installed rigidly fixed to the body and rotated with it as one solid body. We will take as the OX, OY and OZ axes the main central axes of inertia of the fictitious solid body obtained in this way; the moments of inertia of the given body around these axes we will call respectively A, B and C. The first summand in the expression (1) is now written in the simplest form

$$\int_S [\mathbf{r}[\omega\mathbf{r}]]\,dm = Ap\mathbf{e}_x + Bq\mathbf{e}_y + Cr\mathbf{e}_z$$

where $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ are unit vectors parallel to the OX, OY and OZ axes. Using the expressions for projections of the $\boldsymbol{\omega}_i$ vectors on the axes of the $OXYZ$ system we can write the second summand in expression (1) in the form

$$\sum_{i=1}^{n} \boldsymbol{\omega}_i J_i = \mathbf{e}_x \sum_{i=1}^{n} \omega_i a_{ix} J_i + \mathbf{e}_y \sum_{i=1}^{n} \omega_i a_{iy} J_i + \mathbf{e}_z \sum_{i=1}^{n} \omega_i a_{iz} J_i$$

Grouping according to the unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$, we finally obtain for $\mathbf{K}$

$$\mathbf{K} = \left(Ap + \sum_{i=1}^{n} \omega_i a_{ix} J_i\right)\mathbf{e}_x + \left(Bq + \sum_{i=1}^{n} \omega_i a_{iy} J_i\right)\mathbf{e}_y + \left(Cr + \sum_{i=1}^{n} \omega_i a_{iz} J_i\right)\mathbf{e}_z \qquad (2)$$

We will apply the theorem of kinetic moment to the motion of the given system relative to the mass centre; since the mass centre of the system O is taken as the point of reduction of the kinetic moment vector then we can write

$$\frac{d\mathbf{K}}{dt} = \mathbf{M} \qquad (3)$$

Here $d\mathbf{K}/dt$ is the absolute derivative of the $\mathbf{K}$ vector in time, and $\mathbf{M}$ is the principle moment of all the external forces acting on the system. By way of external moments with reference to the system, we have to consider the moment of external perturbations $\mathbf{M}_B$ acting on the missile and the guiding moment of the jet engines $\mathbf{M}_p$ (if, in addition to reactive flywheels of the stabilizing system, the missile employs jet engines by way of supplementary mechanisms),* and

$$\mathbf{M} = \mathbf{M}_B + \mathbf{M}_p$$

Calculating the absolute derivative of the K vector as the sum of the relative and transfer velocities of the end of the K vector in the system $OXYZ$ we obtain

$$\frac{d\mathbf{K}}{dt} = \dot{K}_x\mathbf{e}_x + \dot{K}_y\mathbf{e}_y + \dot{K}_z\mathbf{e}_z + [\omega\mathbf{K}]$$

Here $\dot{K}_x$, $\dot{K}_y$ and $\dot{K}_z$ are time derivatives of the projections of the $\mathbf{K}$ vector on the OX, OY and OZ axes:

$$\dot{K}_x = A\dot{p} + \sum_{i=1}^{n} \dot{\omega}_i a_{ix} J_i$$

$$\dot{K}_y = B\dot{q} + \sum_{i=1}^{n} \dot{\omega}_i a_{iy} J_i$$

$$\dot{K}_z = C\dot{r} + \sum_{i=1}^{n} \dot{\omega}_i a_{iz} J_i$$

Equating the respective projections of the $d\mathbf{K}/dt$ and $\mathbf{M}$ vectors on the axes of the $OXYZ$ system, we obtain (summation indices from 1 to n are omitted in the following):

$$A\dot{p} + \Sigma\dot{\omega}_i a_{ix} J_i + q(Cr + \Sigma\omega_i a_{iz} J_i) - r(Bq + \Sigma\omega_i a_{iy} J_i) = M_{Bx} + M_{px}$$

$$B\dot{q} + \Sigma\dot{\omega}_i a_{iy} J_i + r(Ap + \Sigma\omega_i a_{ix} J_i) - p(Cr + \Sigma\omega_i a_{iz} J_i) = M_{By} + M_{py}$$

$$C\dot{r} + \Sigma\dot{\omega}_i a_{iz} J_i + p(Bq + \Sigma\omega_i a_{iy} J_i) - q(Ap + \Sigma\omega_i a_{ix} J_i) = M_{Bz} + M_{pz}$$

The system obtained is transformed as follows:

$$A\dot{p} - \Sigma a_{ix}(a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r}) J_i + \Sigma a_{ix}(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r}) + qr(C - B) = M_{Bx} + M_{px} - q\Sigma\omega_i a_{iz} J_i + r\Sigma\omega_i a_{iy} J_i$$

* On introducing jet engines it will be necessary to calculate the mass of the given system and all its remaining inertial parameters—moments of inertia, principal axes of inertia, etc.—as variables; however the masses necessary for controlling the position of a missile around its mass centre are small and can be disregarded in calculating the mass variation.

$$B\dot{q} - \Sigma a_{iy}(a_{iz}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + \Sigma a_{iy}(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + rp(A - C) = M_{By} + M_{py} - r\Sigma\omega_i a_{ix}J_i + p\Sigma\omega_i a_{iz}J_i$$

$$C\dot{r} - \Sigma a_{iz}(a_{iz}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + \Sigma a_{iz}(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + pq(B - A) = M_{Bz} + M_{pz} - p\Sigma\omega_i a_{iy}J_i + q\Sigma\omega_i a_{ix}J_i \quad (4)$$

We will clarify the concept of these three summands in the left-hand sides of the equations. We will study the motion of the ith flywheel in the system of co-ordinate axes associated with the flywheel. At the same time as one of the associated axes, we select the support axis of the flywheel with the same reference direction on it as was selected above; the reference origin is located in the mass centre of the flywheel. The two remaining axes are located arbitrarily in a plane perpendicular to the first axis. By reason of the axial symmetry of the flywheel its moments of inertia around these two axes are mutually equal. We will call them J_{ix} (equatorial moment of inertia of the ith flywheel). Since the selected system of axes with reference to the given flywheel represents the system of principal central axes of inertia then we can employ as the equations of motion of the flywheel the dynamic Euler equations.

The projection of the total angular velocity of the flywheel on the first co-ordinate axis (support axis) is determined by the relation

$$p' = \omega_i + a_{ix}p + a_{iy}q + a_{iz}r$$

The projections of the total angular velocity of the flywheel on the two remaining co-ordinate axes will be termed q' and r'. With these notations the first of the dynamic Euler equations (the equation relating to the projections on the first co-ordinate axis) will be written:

$$J_i(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r}) = M'_i$$

Here M'_i is the projection of the total moment applied to the flywheel on the positive direction of its axis. Obviously M'_i is equal in absolute value to the axial moment on the flywheel shaft and carries a positive sign, if the axial moment acting on the flywheel forms a right-hand spiral with positive direction, or negative if the axial moment acts in the opposite direction. The axial moment on the flywheel shaft may be governed only by the active moment of the engine putting the flywheel into rotation $-M'_{ai}$ and the resistance moment on the flywheel shaft M'_{ci}.

$$M'_i = M'_{ai} + M'_{ci}$$

Here M'_{ai} and M'_{ci} are algebraic quantities reckoned exactly in the same way as the quantity M'_i. Then,

$$J_i(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r}) = M'_{ai} + M'_{ci}$$

Multiplying this relation by a_{ix} and summating for i from 1 to n we obtain the expression for the third summand in the first equation of the system (4):

$$\Sigma a_{ix}(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i = \Sigma a_{ix}(M'_{ai} + M'_{ci})$$

Similar expressions can readily be obtained also for the corresponding terms in the remaining two equations

$$\Sigma a_{iy}(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i = \Sigma a_{iy}(M'_{ai} + M'_{ci})$$

$$\Sigma a_{iz}(\dot{\omega}_i + a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i = \Sigma a_{iz}(M'_{ai} + M'_{ci})$$

Substituting for the active moments and the resistance moments at the flywheel shafts M'_{ai} and M'_{ci} applied to the corresponding flywheels the directly opposite moments M_{ai} and M_{ci} applied to the missile body we obtain finally

$$A\dot{p} - \Sigma a_{ix}(a_{ix}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + qr(C - B) = M_{Bx} + M_{px} + \Sigma a_{ix}(M_{ai} + M_{ci}) - q\Sigma\omega_i a_{iz}J_i + r\Sigma\omega_i a_{iy}J_i$$

$$B\dot{q} - \Sigma a_{iy}(a_{iz}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + rp(A - C) = M_{By} + M_{py} + \Sigma a_{iy}(M_{ai} + M_{ci}) - r\Sigma\omega_i a_{ix}J_i + p\Sigma\omega_i a_{iz}J_i$$

$$C\dot{r} - \Sigma a_{iz}(a_{iz}\dot{p} + a_{iy}\dot{q} + a_{iz}\dot{r})J_i + pq(B - A) = M_{Bz} + M_{pz} + \Sigma a_{iz}(M_{ai} + M_{ci}) - p\Sigma\omega_i a_{iy}J_i + q\Sigma\omega_i a_{ix}J_i \quad (5)$$

The last two summands in the right-hand sides of equations (5) represent the projections of the gyroscopic moments applied to the missile body and governed by the forced precession of the flywheels located in the body; the terms

M_{gx}, Σa_{ix}, M_{ai}, etc., yield the projections of the guiding moments, generated by the supplementary organs of the stabilization system; these moments must depend on the angular co-ordinates of the missile and their derivatives, and may also be determined by the special missile position control programme or by control signals from the Earth.

We will now consider the particular case of missile stabilization by means of three reactive flywheels; it is assumed at the same time that the axes of rotation of the first, second and third flywheels are parallel respectively to the OX, OY and OZ axes of the associated $OXYZ$ system, selected within the missile body in the manner indicated above; at the same time the positive directions of the axes of rotation coinside with the positive directions of the corresponding axes of the $OXYZ$ system. Then for the direction cosines we have:

$$\begin{Vmatrix} a_{1x} & a_{1y} & a_{1z} \\ a_{2x} & a_{2y} & a_{2z} \\ a_{3x} & a_{3y} & a_{3z} \end{Vmatrix} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}$$

We will further write

$$\begin{array}{lll} \omega_1=\omega_x & \omega_2=\omega_y & \omega_3=\omega_z \\ M_{a1}=M_{ax} & M_{a2}=M_{ay} & M_{a3}=M_{az} \\ M_{c1}=M_{cx} & M_{c2}=M_{cy} & M_{c3}=M_{cz} \end{array}$$

and assuming identical axial moments of inertia of the flywheels ($J_1=J_2=J_3=J$) we obtain the equations of motion of the stabilized missile in this case

$$\begin{aligned} (A-J)\dot{p}+(C-B)qr &= M_{Bx}+M_{gx}+ \\ &+M_{ax}+M_{cx}-\omega_z qJ+\omega_y rJ \\ (B-J)\dot{q}+(A-C)rp &= M_{By}+M_{gy}+ \\ &+M_{ay}+M_{cy}-\omega_x rJ+\omega_z pJ \\ (C-J)\dot{r}+(B-A)pq &= M_{Bz}+M_{gz}+ \\ &+M_{az}+M_{cz}-\omega_y pJ+\omega_x qJ \end{aligned} \tag{6}$$

The equations obtained are similar to the usual dynamic Euler equations, written for the missile body. Actually the only difference between the left-hand sides of the equations obtained from the left-hand sides of the Euler equations is the presence of the correction J in the first terms of the equations; introducing the same corrections in the second terms of the equations in the following manner

$$(C-B)qr=[(C-J)-(B-J)]qr$$

and so on, it would not be difficult to obtain complete similarity between the left-hand sides of equations (6) and the left-hand sides of the Euler equations. The difference would reside only in the fact that instead of the quantities A, B and C we would have the differences $(A-J)$, $(B-J)$, $(C-J)$. The right-hand sides of equations (6) display complete agreement with the right-hand sides of the Euler equations, written for the missile body; they contain the projections of all the moments occurring with respect to the missile body with external moments. In addition to the moments $\mathbf{M}_B$ and $\mathbf{M}_g$ such moments for the missile body will obviously be the active moments and the resistance moments of the flywheel shafts, transmitted to the body together with the gyroscopic moments of the flywheels, the projections of which $\omega_y rJ$, $\omega_z qJ$, etc., are also contained in the right-hand sides of the equations obtained.

This similarity is, however, superficial, governed by the fact that both the Euler equations and the equations (6) are equations of motion of a system consisting of one or respectively four solid bodies applying to the reference system in which the expressions of the projections of kinetic moment on the co-ordinate axes are the simplest. It should be borne in mind that, as distinct from the case of the Euler equations, the OX, OY and OZ axes, in the projections on which the equations (6) were written, are not, in relation to the missile body, either the principal or the central axes of inertia, and that neither the quantities A, B and C nor the differences $(A-J)$, $(B-J)$ and $(C-J)$ figuring in equations (6) coincide with the moments of inertia of the missile body around these axes.

2. COROLLARIES OF THE LAW OF CONSERVATION OF KINETIC MOMENT

In order to connect the non-rating reference system $OX_0Y_0Z_0$ with the reference system $OXYZ$, associated withe missile body, we will introduce three angles θ, ψ and φ (Fig. 2).

We will define the angle θ as the angle

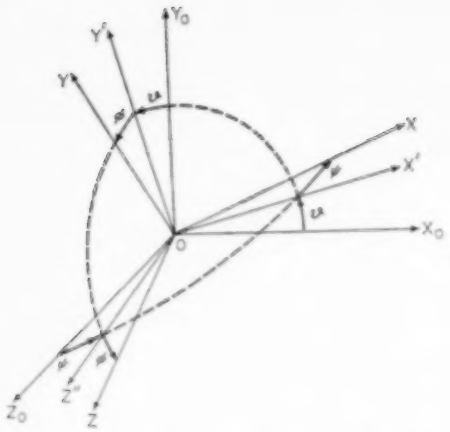


Fig. 2. Sequence of turns at angles θ, ψ, and ϕ for transition of the $OX_0Y_0Z_0$ systems to the OXYZ position.

between OX', the projection of the OX axis on the X_0OY_0 plane, and the OX_0 axis. The angle ψ we will define as the angle between OX' and the OX axis. The angle ϕ we will define as the angle between the OY axis and the line of intersection of the X_0OY_0 and YOZ planes. In the given case we will adopt the following sequence of missile body turns for transition from the position where the $OXYZ$ and $OX_0Y_0Z_0$ systems coincide to the given position of the $OXYZ$ system. The first turn is around the OZ_0 axis at angle θ, then a turn round the new position of the OY_0 axis (OY' in Fig. 2) at angle ψ and finally a turn at angle ϕ round the OX axis. We will draw up a table of direction cosines for the axes associated with the missile and the non-rotating systems of co-ordinates.

We will express the projections of angular velocity of the missile body by the angles θ, ψ and ϕ.

$$p=\dot{\phi}-\dot{\theta}\sin\psi$$
$$q=\dot{\psi}\cos\phi+\dot{\theta}\sin\phi\cos\psi$$
$$r=\dot{\theta}\cos\phi\cos\psi-\dot{\psi}\sin\phi \qquad (7)$$

The motion of the missile relative to its mass centre is expressed by the equation (3). If no external moments are applied to the missile ($\mathbf{M}=0$), then from (3) we obtain the abvious corollary known in mechanics as the law of conservation of kinetic moment

$$\mathbf{M}=\mathbf{C}$$

where $\mathbf{C}$ is some constant vector.

On the basis of this law, a number of interesting conclusions can be reached, regarding the problem of missile rotations around its mass centre.

Let the projections of the $\mathbf{K}$ vector on the axis of the non-rotating system of co-ordinates in inertial space be C_{x0}, C_{y0}, C_{z0}. Then its projection in the system of co-ordinates $OXYZ$ will be connected with C_{x0}, C_{y0}, C_{z0} by the following relations:

$$C_{x0}=K_x\cos\psi\cos\theta+K_y(\sin\phi\sin\phi\cos\theta-$$
$$-\cos\phi\sin\theta)+K_z(\cos\phi\sin\psi\cos\theta+\sin\phi\sin\theta)$$
$$C_{y0}=K_x\cos\psi\sin\theta+K_y(\sin\phi\sin\psi\sin\theta+$$
$$+\cos\phi\cos\theta)+K_z(\cos\phi\sin\psi\sin\theta-\sin\phi\cos\theta)$$
$$C_{z0}=-K_x\sin\psi+K_y\sin\phi\cos\psi+K_z\cos\phi\cos\psi \qquad (8)$$

In addition the apparent equality is correct

$$C_{x0}^2+C_{y0}^2+C_{z0}^2=K_x^2+K_y^2+K_z^2$$

The values K_x, K_y and K_z can readily be obtained from formula (2).

Table 1.

Axes of the non-rotating system	Axes of the associated system		
	X	Y	Z
X_0	$\cos\psi\times\cos\theta$	$-\cos\phi\times\sin\theta+\sin\phi\times\sin\psi\times\cos\theta$	$\cos\phi\times\sin\psi\times\cos\theta+\sin\phi\times\sin\theta$
Y_0	$\cos\psi\times\sin\theta$	$\cos\phi\times\cos\theta+\sin\phi\times\sin\psi\times\sin\theta$	$\cos\phi\times\sin\psi\times\sin\theta-\sin\phi\times\cos\theta$
Z_0	$-\sin\psi$	$\sin\phi\times\cos\psi$	$\cos\phi\times\cos\psi$

We will consider the motion of a missile characterized by the fact that the control system installed on it maintains a given plane (associated with the missile) by means of flywheels parallel throughout to a given plane in stationary space. Without destroying the generality, we can assume that the OX and OY axes of the stationary system of co-ordinates lie in the plane indicated parallel to the OX_0Y_0 plane. Then in undisturbed motion $\psi=\phi=0$ and the equations (8) take the form:

$$\begin{aligned} C_{x0} &= K_x \cos\theta - K_y \sin\theta \\ C_{y0} &= K_x \sin\theta + K_y \cos\theta \\ C_{z0} &= K_z \end{aligned} \tag{9}$$

In the reduced relations (9) the angle θ may be a given function of time. Assume for example that the missile passes around some celestial body along the orbital plane and that it is required from the control system that the axis associated with the missile lying in the XOY plane should be directed permanently towards the centre of the celestial body whilst the XOY plane itself should be parallel to the orbital plane. Then during the period of one circle the angle θ varies by 2π since the missile body performs one turn around the OZ axis.

The first two equations (9) show that with such a motion around a celestial body the values K_x and K_y will be periodic functions of time in the same way as the angle θ.

It can readily be seen that

$$\begin{aligned} K_x &= C_{xy} \sin[\theta(t)+\theta] \\ K_y &= C_{xy} \cos[\theta(t)+\theta] \end{aligned} \tag{10}$$

where

$$C_{xy} = \sqrt{C_{x0}^2 + C_{y0}^2} \text{ and } \theta = \operatorname{arctg}\frac{C_{x0}}{C_{y0}}$$

For disturbed motion, as at present being studied, the angles $\psi=\phi=0$ and the angular velocities of the missile body, as seen from formulae (7) will be

$$p=0,\ q=0,\ r=\dot{\theta}$$

In this case, following from formula (2), the projections of the kinetic moment K_x and K_y different from zero may be connected only with the rotation of flywheels which achieve certain angular velocities over the period preceding the given moment. In the given case it is thus possible to consider K_x and K_y as projections of the kinetic moments of the flywheels and consequently the number of revolutions of the flywheels, the axes of which are stationary relative to the missile body, must generally speaking vary periodically. In the simplest case, when three flywheels are installed in the missile, the axes of rotation of which are parallel to the OX, OY and OZ axes, the periodically varying number of revolutions will apply to the flywheels whose axes are parallel to OX and OY. In accordance with the notations adopted earlier in the case where projections of the kinetic moment of the missile body on the OX and OY axes are equal to zero we obtain

$$K_x = \omega J_x \text{ and } K_y = \omega_y J_y \tag{11}$$

where J_x and J_y are the moments of inertia of the flywheels installed parallel to the OX and OY axes respectively whilst ω_x and ω_y are their angular velocities (the i and ω indexes are omitted since in the given case no confusion can arise).

It follows from a comparison of (10) and (11) that ω_x and ω_y are periodic functions of time. During the period of one turn of the missile around the central body the angular velocity of each flywheel attains a maximum value, falls to zero, changes its sign and having reached the same maximum value in absolute terms again falls to zero and returns again to the initial value. This behaviour of the flywheels can be understood if we consider that for $r\neq 0$ the angular velocities of the flywheels ω_x and ω_y yield gyroscopic moments differing from zero in the first two equations (6); these moments may be damped by the corresponding boost or braking of the flywheels. Damping of the gyroscopic moments is necessary for satisfying the requirement $p=0$ and $q=0$.

It is important to observe that boosting and braking, inevitably involving the expenditure of energy, are by no means essential for uniform motion along the orbit. Indeed, if at the initial moment the projections of the **K** vector on the orbital plane were equal to zero, that is to say $C_{x0}=C_{y0}=0$ then as seen from the relations (10) and (11), $\omega_x=\omega_y=0$ throughout the entire period of motion along the given orbit. Hence

if the effort is towards reducing the expenditure of energy, then the reactive flywheels whose axes lie on the orbital plane should be braked (for example, by employing the moments set up by jet engines).

Braking of the flywheels is by no means always possible, and for this reason a somewhat closer study is necessary of the problem of energy expenditures on periodic development of the flywheels.

The kinetic energy of flywheels is expressed by the following equations

$$\frac{1}{2}J_x\omega_x^2=\frac{1}{2J_x}C_{xy}^2\sin^2[\theta(t)+\theta]$$

$$\frac{1}{2}J_y\omega_y^2=\frac{1}{2J_y}C_{xy}^2\cos^2[\theta(t)+\theta]$$

Adding these two equations term by term it can readily be seen that for $J_x\neq J_y$ the total kinetic energy of the flywheels is a periodic function of time. Consequently, in the general case, part of the mechanical energy of the flywheels must be transformed periodically into some other form (for example, electrical) and then again be transformed into energy of rotation of the flywheels. It should be observed that for $J_x=J_y$ also, the braking of one and simultaneous development of the other flywheel should practically always be carried out by conversion of the kinetic energy of the braked flywheel into another form of energy and development of the second flywheel by some motor, driven by non-mechanical energy. In this way, in the general case, operation of the flywheels is inevitably associated with the conversion of energy which in turn involves additional losses. We will estimate these losses.

Let the efficiency of conversion of mechanical energy into some other form, allowing reverse energy conversion, be η_1 and the efficiency of the reverse energy conversion η_2. The quantity of mechanical energy passing into the given form with complete braking of the flywheel will be equal to $1/2J\eta_1C_{xy}^2$, whilst the irreversible losses of mechanical energy are $1/2J(1-\eta_1)C_{xy}^2$. With acceleration of the flywheel there is expended the energy $1/2J\eta_2C_{xy}^2$, which exceeds the kinetic energy obtained by the quantity $(1/\eta_2-1)1/2JC_{xy}^2$. Hence the total energy loss during the time from braking the flywheel up to acceleration will be

$$\Delta E=\frac{1}{2J}\left(\frac{1}{\eta_2}-\eta_1\right)C_{xy}^2$$

The equation obtained shows that for given values C_{xy}, η_1 and η_2 a reduction in the losses can be achieved only by raising the moment of inertia of the flywheel.

An increase in J however involves an increase in the dimensions or mass of the flywheel or both simultaneously. We will therefore examine the problem of optimum values for J for two flywheels on the condition that the total mass of the flywheels is limited. The total energy losses on the two flywheels for one braking and acceleration cycle will be

$$\Delta E=\frac{1}{2}\left(\frac{1}{J_x}+\frac{1}{J_y}\right)\left(\frac{1}{\eta_2}-\eta_1\right)C_{xy}^2$$

In the case of a given design solution for a flywheel its mass m may usually be represented approximately in the form of a power function of inertia $m=aJ^n$, where a is a constant, and $n>0$. The condition $n>0$ is quite natural since with increasing J the flywheel mass also increases.

The condition of constant total mass of the two flywheels is expressed by the equation

$$J_x^n+J_y^n=b \tag{12}$$

where b is a constant value.

We will seek the minimum ΔE observing the condition (12). Employing the Lagrange method we will seek the extremum of the function Z of the two variables J_x and J_y

$$Z=\frac{1}{2}\left(\frac{1}{J_x}+\frac{1}{J_y}\right)\left(\frac{1}{\eta_2}-\eta_1\right)C_{xy}^2+\lambda(J_x^n+J_y^n-b)$$

where λ is the Lagrange multiple. After elementary calculations we obtain the extremum condition $J_x=J_y$. It can readily be demonstrated that this is the minimum condition.

Hence for reducing the energy loss in the control system it is necessary as far as possible to brake the flywheels, and if this is not possible, then to increase their moments of inertia whilst the moments of inertia of the fly-

wheels in mutual action along the orbital path must be chosen identical.

3. EXAMPLE OF APPLICATION OF THE OBTAINED RELATIONS TO THE CASE OF SMALL DEVIATIONS OF THE MISSILE FROM A GIVEN POSITION

By way of example we will take the case of stabilization of a missile around a given invariable position in space. Let the task of the stabilization system be the maintenance of a missile position such that the associated system of co-ordinate $OXYZ$ coincides throughout with some non-rotating system of co-ordinates $OX_0Y_0Z_0$.

As before, the mutual arrangement of the $OXYZ$ and $OX_0Y_0Z_0$ systems will be characterized by the angles θ, ψ and ϕ introduced above. With regard to the orientation of the flywheel axes in the missile body and the values of the axial moments of inertia of the flywheels we will adopt the same assumptions as were adopted in drawing up equations (6). The value of the axial moment of inertia of the flywheels J as compared with the moments of inertia A, B and C will be disregarded. We will further assume $\mathbf{M}_B = \mathbf{M}_p = 0$ and we will disregard the values of the gyroscopic moments and the resistance moments as compared with the values of the control moments M_{ax}, M_{ay}, M_{az}. We will confine ourselves to the case of small deviations of the missile from the given position. The expressions (7) are then simplified as follows:

$$p = \dot{\phi},\ q = \dot{\psi},\ r = \dot{\theta}$$

If it is assumed further that the derivatives of the projections of the angular velocity $\dot{p}$, $\dot{q}$, $\dot{r}$ have the same order of smallness as the projections p, q, r themselves, then the second summands of the left-hand sides of equations (6), as terms of the highest order of smallness, can be discarded and the equations take the form

$$A\ddot{\phi} = M_{ax}$$
$$B\ddot{\psi} = M_{ay}$$
$$C\ddot{\theta} = M_{az}$$

Assume that the moment on the shaft of the reactive flywheel established along the OX axis is a function of the angular velocity developed by it and depends further (through some control system) on the angle ϕ and its derivative $\dot{\phi}$

$$M_{ax} = -M'_{ax}(\omega_x,\ \phi,\ \dot{\phi})$$

If similar equations can be written for M_{ay} and M_{az} then we obtain

$$A\ddot{\phi} = -M'_{ax}(\omega_x,\ \phi,\ \dot{\phi})$$
$$B\ddot{\psi} = -M'_{ay}(\omega_y,\ \psi,\ \dot{\psi})$$
$$C\ddot{\theta} = -M'_{az}(\omega_z,\ \theta,\ \dot{\theta})$$

It thus follows that the missile oscillations around the OX, OY and OZ axes are independent.

We will study more closely the first of the equations. We will decompose the function M'_{ax} into a Taylor series near the point $(\omega_{x0},\ 0,\ 0)$ corresponding to the undisturbed conditions. Maintaining the first terms of the decomposition we obtain

$$M'_{ax}(\omega_x,\ \phi,\ \dot{\phi}) = M'_{ax}(\omega_{x0},\ 0,\ 0) + \frac{\partial M'_{ax}}{\partial \omega}\delta\omega_x + \frac{\partial M'_{ax}}{\partial \phi}\phi + \frac{\partial M'_{ax}}{\partial \dot{\phi}}\dot{\phi}$$

where partial derivatives are taken at the point $(\omega_{x0},\ 0,\ 0)$. The motion of the missile in the YOZ plane will be described by the equation

$$A\ddot{\phi} + \frac{\partial M'_{ax}}{\partial \dot{\phi}}\dot{\phi} + \frac{\partial M'_{ax}}{\partial \phi}\phi + \frac{\partial M'_{ax}}{\partial \omega}\delta\omega_x = M_{ax}(\omega_{x0},\ 0,\ 0)$$

In order to remove $\delta\omega_x$, use is made of the first equation (8). For the case of small angles θ, ψ, ϕ, discarding the terms of the first and higher orders of smallness we obtain $C_{x0} = K_x$ or $C_{x0} = A\dot{\phi} + J\omega_x$. Varying in the neighbourhood of the given point we find

$$\delta\omega_x = -\frac{A}{J}\dot{\phi}$$

Consequently the final equation of motion

of the missile in the YOZ plane will take the form:

$$A\ddot{\phi} + \left(\frac{\partial M'_{ax}}{\partial \dot{\phi}} - \frac{A}{J}\frac{\partial M'_{ax}}{\partial \omega_x}\right)\dot{\phi} + \frac{\partial M'_{ax}}{\partial \phi}\phi = M_{ax}(\omega_{x0}, 0, 0)$$

The term $M_{ax}(\omega_{x0}, 0, 0)$ in the right-hand side of the equation represents the displacement effect governing the static error of the system

$$\phi_{stat} = \frac{M_{ax}(\omega_{x0}, 0, 0)}{\dfrac{\partial M'_{ax}}{\partial \phi}}$$

It is interesting to note that the dependence of the moment M_{ax} on the angular velocity of the flywheel for $\delta M'_{ax}/\delta\phi < 0$ assists damping of the perturbations. With a sufficiently close relationship of the moment M_{ax} to the angular velocity ω_x it is possible not to introduce the derivative of the angle into the law of regulation.

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 107–116. Printed in Great Britain.

ORBITS OF SPACE ROCKETS IN THE DIRECTION OF THE MOON

L. I. SEDOV*

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli* No. 5, p. 3 (1960)

The performance and utilization of interplanetary flights is based on a theoretical analysis of the laws of motion of vehicles in cosmic space and on numerical calculations carried out on the basis of these relationships. The main requirements relating to the energetics and control programme of the motion of rocket carriers together with the optimum and acceptable launching conditions are established by calculation.

By way of an initial system of reckoning for describing the path of vehicles a system of Cartesian co-ordinates can be taken with origin at the Earth's centre, moving progressively relative to the stars. In the case of several practical problems it is necessary to use spherical systems of co-ordinates related invariably with the Earth and having the origin of the co-ordinates at the Earth's centre or at different points on its surface. It is sometimes necessary to consider the path or object relative to the Moon or other planets.

Into the differential equations of motion of celestial ballistics describing free flight in cosmic space it is necessary to introduce only the forces of interaction, governed by Newton's law of universal attraction. In solving problems of alighting on the Moon and circumnavigating the Moon it is necessary to examine the path of a body in a known gravitational field governed by the Sun, the Moon and Earth (taking account of the Earth's oblateness).

In order to probe the optimum conditions for launching rockets to the Moon use can be made of approximate methods and the study of the path of a body as Kepler motion relative to Earth, when the distance of the Moon is greater than 66,000 km, and as Kepler motion relative to the Moon when the distance is less than 66,000 km.

For given points in the northern hemisphere of the Earth the optimum conditions for launching can be obtained to a first approximation (taking account only of the Earth's gravity) employing the following concepts. We will examine the impact version. Call A the position of the end of the active section of the rocket, B the position of the centre of the Moon at the moment of impact and O the Earth's centre (Fig. 1). Each orbit corresponding to impact from point A to point B lies on a plane governed by three points A, O, B. For the complete determination of the orbit and of the initial velocity v it is sufficient to apply an angle of inclination θ to the horizon of the initial velocity v.

Resolving the problem of two bodies, the functional relationship can readily be found

$$\frac{v^2}{v_p^2}=\gamma\left(\frac{r}{R},\ \theta, \Phi\right) \tag{1}$$

where v_p is the parabolic velocity. The angle between the directions OA and OB, termed Φ, is called the angular range. The point A, connected invariably with the Earth, describes a parallel with the rotation of the Earth. At the same time the angle Φ varies over a range which can be readily found from Fig. 1.

Since the relationship $\frac{r}{R}$ is small and practically constant, then according to (1) the relationship $\frac{v^2}{v_p^2}$ is basically a function of θ and Φ.

In Fig. 2 this relationship of θ for $\Phi=$const. is indicated graphically with the continuous lines. The discontinuous lines on the same graph show a typical form of rocket characteristics with various constant values for the

* The article is the text of a survey paper presented at the Annual Meeting of the American Rocket Society in Nov. 1959.

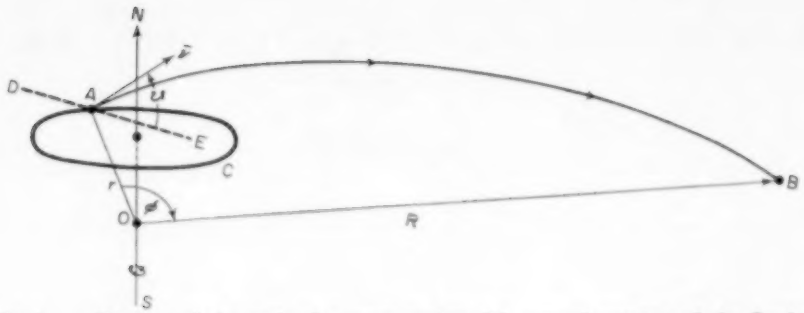


Fig. 1. Diagram of the path of a space rocket. *SN*—axis of rotation of the Earth; *AC*—parallel containing the point *A*; *DE*—line of the horizon in the orbital plane.

weight of the rocket at the end of the boost phase. It follows from Fig. 2 that the optimum launching moment for which the effective weight is the greatest, corresponds to the maximum range angle Φ. It is apparent that in this case the orbital plane is perpendicular to the equatorial plane. The maximum angle Φ indicated depends on the position of the Moon in its orbit around the Earth. It is clear that for points in the northern hemisphere the largest value Φ corresponds to the lowest position of the Moon under the equatorial plane. For the given point A these conditions determine the most favourable launching time during the lunar month and the launching moment during the Earth day.

These results were obtained disregarding the rotation of the Earth and the suitability of using the additional velocity component of the rocket obtaining on this account. Taking account of the rotation of the Earth reduces the inclination of the optimum orbit to the equatorial plane. It is necessary also to introduce corrections connected with the influence of the Moon, the Sun, the Earth's oblateness and the properties of the acceleration boost

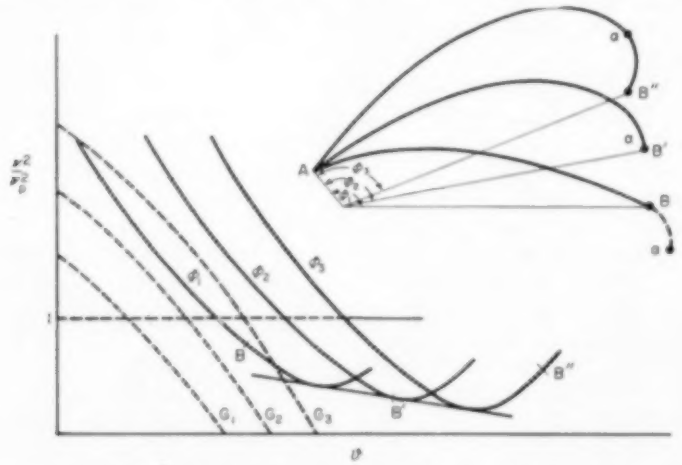


Fig. 2. $\frac{v^2}{v_p^2}$ as a function of the angle of inclination θ for various values Φ (continuous curves: $\Phi_1 > \Phi_2 > \Phi_3$) and for various weights of rocket (discontinuous curves: $G_1 < G_2 < G_3$). The points B, B' and B'' correspond to impact of the missile on the Moon. On the right are shown diagrams of the correspondiing trajectories; a—position of apogee, A—escape point on the trajectory.

effective weight may be obtained by the appro-phase. A further substantial increase in the effective weight may be obtained by the appropriate selection of the launching point on the Earth's surface. The most favourable points for launching are near the Equator.

For the Soviet space rockets a flight was selected with the plane of the initial section of the orbit inclined at 65° to the equatorial plane equal, which approaches the most favourable. The main calculations of the laws of motion were carried out on high-speed computers. Optimum trajectories were found and a study was made of families of trajectories near to optimum.

Selection of a given orbit can commence by fixing the total specific energy of the launched object at the end of the boost phase. The total energy is determined by the flight problems and the necessity for obtaining favourable conditions for observation of the object from the territory of the U.S.S.R. at the moment of impact or on entering the vicinity of the Moon. It is also governed by the flight time to the Moon. An analysis of optimum conditions shows that the flight time from the Earth to the region of the Moon must be near to 0·5, 1·5, 2·5 or 3·5 days.

In order to carry out launching, data is necessary (after determining the total energy) regarding the possible effective weights and the necessary launching conditions during the course of a few days near to the most suitable day, corresponding for northern latitudes to the smallest lunar declination at the moment of approaching the Moon. This problem is resolved by ballistic calculations taking into account the rocket characteristics.

Calculations show that, with deviations in the launching time from the optimum by 1 or 2 days on both sides, the losses in effective weight are within permissible limits. Therefore it is advisable to carry out launching only during a few days in each lunar month. A substantial deviation in the launching day from the optimum involves substantial losses in the possible effective weight.

For detailed investigations of the influence of the initial parameters on the trajectory characteristics it was found convenient to establish a number of important properties of the family of trajectories with a central trajectory striking the centre of the Moon. The properties listed below of such a family of trajectories are achieved with sufficient accuracy; several practical problems can be resolved very satisfactorily using these properties.

1. Near to the Moon in close proximity around the central trajectory of the type selected, which can be considered as a straight line, near trajectories at distances up to 20,000 km from the Moon form a set of lines possessing rotational symmetry relative to the central axial trajectory.

2. In the vicinity of the Moon the trajectories approach conical sections, lying in the meridianal planes.

3. Different trajectories of this family are obtained with small deviations from the main values of the six independent parameters. For these six parameters the rocket co-ordinates and the absolute velocity components at the end of the acceleration stage can be taken.

4. On varying only any one parameter or with proportional variation in any set of 6 parameters, different trajectories of the group are obtained lying in one and the same meridianal plane.

5. A universal dimensionless variable λ can be introduced into the meridianal planes, being one and the same for all planes governed by the relationship

$$\lambda = \frac{\Delta\xi}{\Delta\xi^*},$$

where $\Delta\xi$ is the increase in parameter ξ at the end of acceleration and $\Delta\xi^*$ is the increase corresponding to the given minimum distance of the disturbed trajectory from the centre of the Moon.

Consequently in any meridianal plane there is a universal relationship for different trajectories:

$$\rho_{\min} = f(\lambda),$$

where $\rho_{\min}$ is the minimum distance of the trajectory from the centre of the Moon.

On the basis of the above properties the six-parameter group is reduced to a two-parameter

group, which can be reduced simply to an arbitrary plane π in which the points are determined by polar co-ordinates: radius λ and polar angle ω (ω signifies the angle governing the meridianal plane in the trajectory space). This circumstance affords a very clear method for investigating and describing the set of disturbed motions.

Isolines of the different numerical characteristics of the trajectories can conveniently be constructed in the π plane. This is particularly important for investigating the round flight trajectory and for finding the initial data corresponding to orbits with the necessary and most suitable characteristics. An example of two families of isolines used for selecting the round flight trajectory of an automatic interplanetary station is shown in Fig. 3.

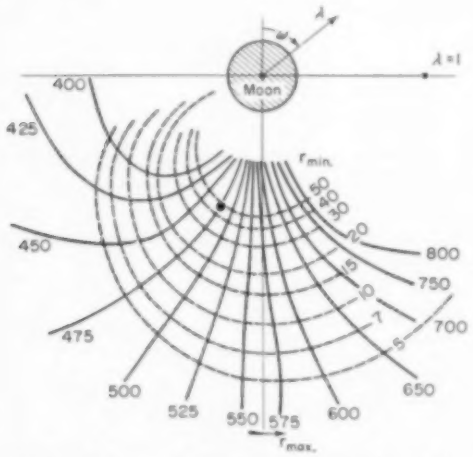


Fig. 3. Isolines of maximum distance from the Earth's centre (continuous curves) and minimum distance (discontinuous curves) in the planes λ, ω. The circle corresponds to the automatic interplanetary station. The figures near the curves relate to the corresponding constant value $r(10^3$ km) for the given curve.

This graph carries families of isolines of maximum and minimum distances from the object to Earth during the first loop after approaching the Moon. For the selected group of trajectories the initial energy (in the absence of influence from the Moon) corresponds to a distance of apogee from the Earth's centre equal to 550,000 km. For trajectories of the group to the left of the "apogee at 550,000 km" isoline there is deceleration and the height of apogee falls off; for trajectories to the right of this isoline there is acceleration, the height of apogee increases and the energy of the object relative to Earth after approaching the Moon increases. The corresponding quantitative effects are clearly visible from the graph.

The values $\rho_{min} = 7{,}900$ and the values λ and ω obtained for the path of an automatic interplanetary station are determined by calculation. After plotting these data on the graph (Fig. 3) we obtain a distance of apogee from the Earth's centre equal to 480,000 km and a distance to perigee on the first loop equal to 47·5 thousand km.

The perturbing effect of the Moon on the path of an automatic station is very substantial; it can be characterized quantitatively by the data of Table I. It is observed that the moment vector of the quantity of motion increases approximately 3 times whilst the direction changes almost oppositely.

A study of the π plane simplifies the investigation of the influence of dispersion of the initial parameters on the orbital characteristics. Introduction of the π plane affords combination of the problem of impact with the Moon and the problems connected with flights passing close to the Moon. Obtaining round flight trajectories with given characteristics is reduced to the problem of impact in the π plane at a given point corresponding to the sought orbital characteristics.

Flight around the Moon can occur along different types of trajectory.

The long flight trajectory (at distances in the order of 40-100,000 km from the Moon, with only slight lunar influence) approaches an ellipse focused on the centre of the Earth. With this type of trajectory and with the launching point in the northern hemisphere the return path (following the approach to the Moon) will pass below the Equator on approaching the Earth. This renders observation of the object from the territory of the U.S.S.R. impossible at close distances from the Earth. In addition, for

Table 1.

	Total geocentric energy $km^2\ sec^{-2}$	Geocentric moment of the quantity of motion. $10^3\ km^2\ sec^{-1}$			
		Cx	Cy	Cz	C
Before the sphere of lunar influence	-0·68	55	-30	21	67
After passing through the sphere of lunar influence	-0·74	-177	65	32	191

Note: Cx is the projection on the vernal equinox axis x; Cy is the projection on an axis perpendicular to the x axis in the equatorial plane; Cz is the projection on the axis of Earth's rotation; $C^2=C_x^2+C_y^2+C_z^2$.

these trajectories the distance of perigee from the Earth's centre is less than the radius of the Earth and therefore the object enters the Earth's atmosphere on the first loop. Consequently the life will be short.

Investigation has shown that another type of trajectory is the most favourable, the path of which takes the object to a distance in the order of 5-20,000 km from the Moon. In this case the Moon presents a strong disturbance as the result of which trajectories can be obtained returning to the vicinity of the Earth, after which the object may convert into a satellite with a high perigee in the order of 40,000 km and with apogee in the order of 500,000 km. In this case the return to Earth after proximity with the Moon will occur in high latitudes of the northern hemisphere of the Earth which offers favourable opportunities for observations and for receiving radio information. In addition these trajectories are very suitable for solving problems regarding the orientation of the object and for photographing from distances in the 40-150,000 km range. Calculations have shown that for this range of distances, traversed in a period of approximately half a day, the direction of the vector "object-Moon" in absolute space is practically maintained. This is suitable for normal operation of the orientation system and photography. This feature of the orbit permits switching of the orientation system and subsequent operations connected with photography by means of a programming unit operating to a previously set time programme.

For the purpose of photographing the invisible side of the Moon it is necessary to carry out launching on days near to new Moon. In this case the invisible side of the Moon is lit by the Sun. The combination of conditions providing suitable energetics with illumination conditions of the Moon is governed by the months and days most suitable for obtaining photographs of the far side of the Moon.

The concepts set out above served as the basis for selecting the round trajectory for the third space rocket whilst for the best solution of the problem use was made of the powerful influence of the Moon on the path of the rocket.

Practical data are set out below regarding the orbits of the three Soviet space rockets.

1. The final stage of the first space rocket without fuel weighed 1472 kg; the weight of the container, scientific equipment and supply sources came to 361·3 kg. The rocket was launched on the 2nd January 1959 towards the Moon; it passed north of the Moon at a distance approximately 5,000 km from its surface. The flight time from Earth to the vicinity of the Moon was 1·5 days.

In launching this rocket, for the first time parabolic velocity was exceeded as a result of which after passing near to the Moon the rocket continued to travel away from Earth, passed out of the sphere of the Earth's gravity and became an artificial solar satellite.

This rocket moves relative to the Sun along an orbit with the following characteristics:

(*a*) the inclination of the orbital plane to the ecliptic plane $i=1°$;

(*b*) minimum distance from the Sun 146×10^6 km;

(*c*) maximum distance from the Sun 197×10^6 km;

(*d*) circling time $T = 450$ days.

The minimum distance from the Sun was reached in the middle of January 1959. The shortest distance of the orbit of the artificial planet to the orbit of Mars is equal to 15×10^6 km.

Radio communication with the first space rocket was maintained up to distances of 400-500,000 km.

2. The second space rocket was launched as an impact version. The weight of the final stage without fuel was 1511 kg, the weight of the container with scientific equipment was 390 kg. A diagram of the rocket path is shown in Fig. 4.

The following data characterize the flight orbit of this rocket. The trajectory lies practically in one and the same plane forming an angle of 65° with the equator. The flight from Earth to the Moon occupied approximately 1·5 days.

The initial velocity in free flight exceeded the local parabolic velocity. From the standpoint of accuracy requirements of the initial data the flight time approximating to 1·5 days was more suitable for the impact version than longer flight times (2·5 days or 3·5 days), for which smaller initial energy is required. The flight to the Moon occupying approximately 0·5 days demands too large an energy.

The minimum geocentric speed in orbit was slightly in excess of 2 km/sec. The velocity of impact with the Moon's surface was 3·3 km/sec. The velocity vector made an angle of approximately 60° with the Moon's surface.

The point of impact with the Moon was 800 km to the north from the centre of the visible disc and south of the Archimedes, Aristillus and Autolycus craters (Fig. 5). The corresponding $\rho_{\min}$ for the continued trajectory is 500 km.

The moment of impact was determined by the sharp cessation of radio signal reception on 14th September at 0 hr 2 min 24 sec universal time. This time agrees with the data of trajectory measurements, carried out both during the flight and during the period immediately preceding the moment of impact.

3. The third space rocket was launched 4th October 1959 in the round-flight version. The weight of the final stage without fuel was 1553 kg. The weight of the automatic interplanetary station together with apparatus and supply sources, installed on the last stage, amounted to 435 kg.

During the flight of the automatic interplanetary station from Earth to Moon the orbital inclination to the equatorial plane was 65°. After perturbation by the Moon the further path under influence of the Earth's

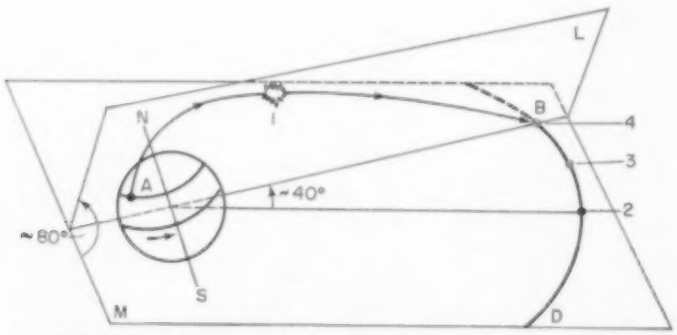


Fig. 4. Diagram of the flight of the second space rocket. *M*—orbital plane of the Moon; *DB*—orbit of the Moon; *L*—plane of the rocket trajectory; *AB*—rocket trajectory; 1—formation of artificial comet; 2—point on the lunar orbit with smallest declination; 3—Moon at moment of launching the space rocket; 4—Moon at moment of impact with rocket.

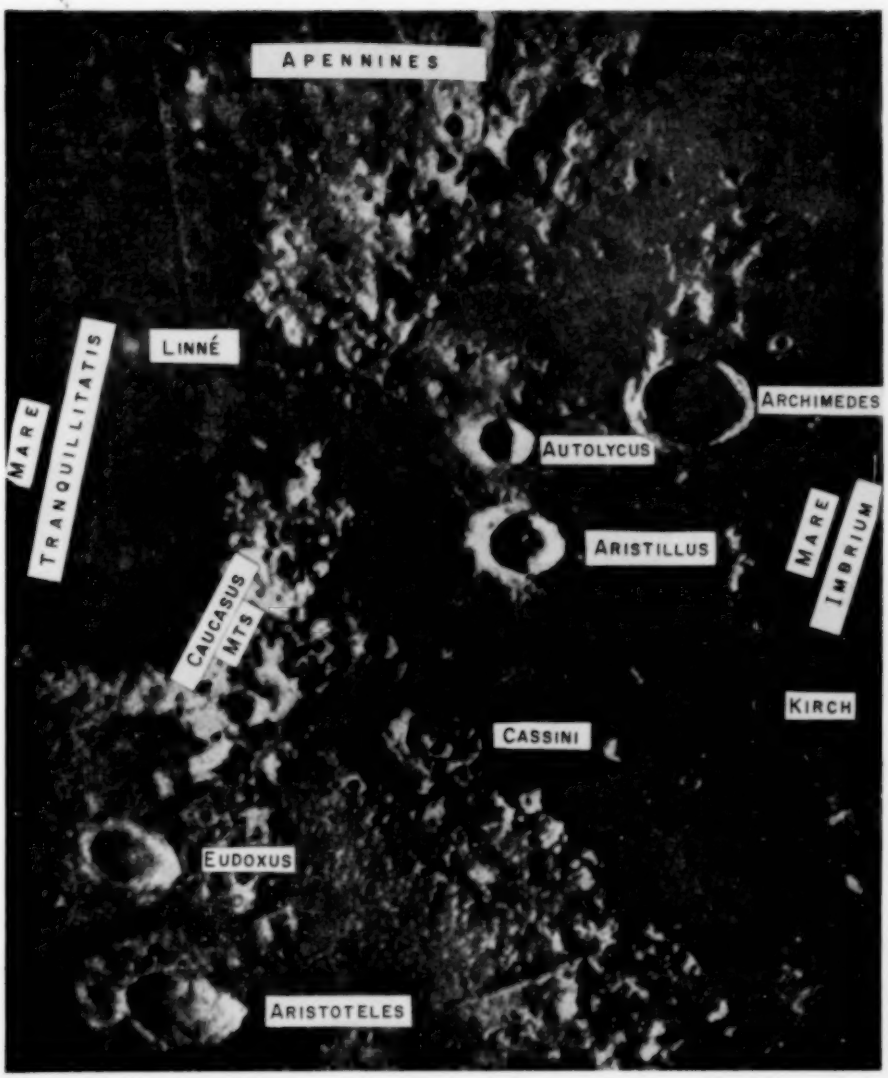


Fig. 5. Section of the Moon's surface containing the site of impact of the second space rocket with the Moon.

(*To face page* 112)

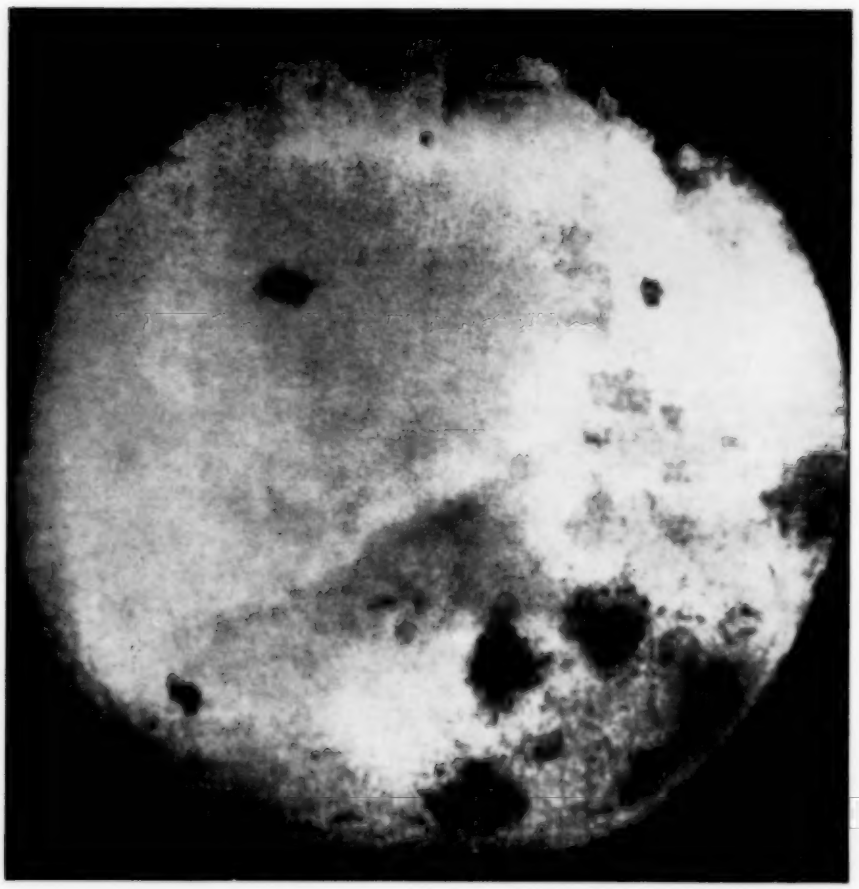


Fig. 10. Photograph of the Moon's surface obtained from the automatic interplanetary station.

gravity followed on orbit near to elliptical with an inclination to the equator approximately 80°. Calculation of the further path shows that the Sun and Moon influence the orbit of the automatic interplanetary station in such a way that the orbital inclination varies irregularly and gradually diminishes. At the tenth loop the inclination is 48°. At the eleventh loop, under the influence of the Moon the orbital inclination increases again to 57°. It is a striking fact that the minimum distance from the Earth as a result of the influence of the Sun and Moon falls off from loop to loop. Calculation shows that after completing the eleventh revolution at the end of March 1960 the automatic interplanetary station entered the Earth's atmosphere in the northern hemisphere and terminated its existence.

This fact is connected with the shape of the orbit and the nature of its situation relative to the Earth and the Sun. This effect, unexpected at first sight, is subject only to Newtonian forces. It is obvious that similar effects must be taken into account in the theoretical analysis of problems of the structure of planetary systems and of orbital characteristics of different planets and their satellites in the solar system. As a result of perturbations set up by the Sun there is an evolution of the orbit which may lead to collision of the satellite with the parent planet; hence only satellites with certain specific types of orbit may persist over a prolonged period.

Fig. 6 shows a projection of the first loop of the orbit of the artificial interplanetary station on the equatorial plane and Fig. 7 on a plane perpendicular to a direction from the Earth's centre to the point of the Vernal equinox.

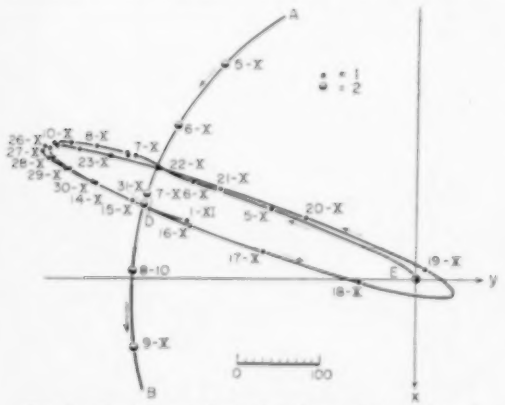


Fig. 6. Projection of the first loop of the orbit of the automatic interplanetary station on the Earth's equatorial plane. *E*—Earth; *AB*—orbit of the Moon; *x* axis directed to the point of Vernal equinox; at point *D* photography of the far side of the Moon was carried out; the points 1 and 2 represent the respective positions of the automatic interplanetary station and the Moon at 0 hrs universal time (3 hrs Moscow time) on the dates indicated in the diagram; the scale of distances is shown below (10^3 km).

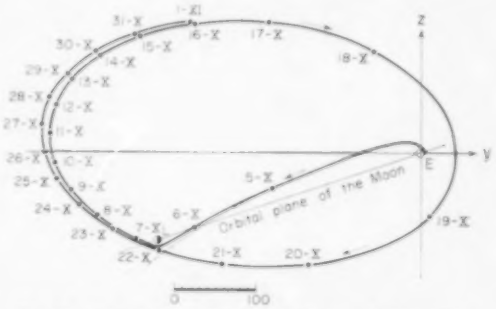


Fig. 7. Projection of the first loop of the automatic interplanetary station orbit on a plane perpendicular to the direction of the Earth's centre—point of Vernal equinox. The *z* axis coincides with the axis of rotation of the Earth; the *y* axis represents the intersection of the equatorial plane with the plane of projection; *L*—Moon; the line *LZ* is the intersection of the orbital plane of the Moon with the plane of projection; the remaining characters are the same as for Fig. 6.

Fig. 8 shows a projection of one loop of the orbit on the Earth's surface. The projection line describes loops at the moment of passing near to the axis of the Earth's rotation both on approaching the Earth and on moving away. Table 2 compares data of characteristic points of the orbit of the automatic interplanetary station.

Fig. 9 shows a projection of the orbit on the surface of the Moon, at the same time indicating points corresponding to the minimum distance to the centre of the Moon and the moment of photographing.

The conditions were favourable for photographing from the standpoint of illumination of the invisible side of the Moon, approximately

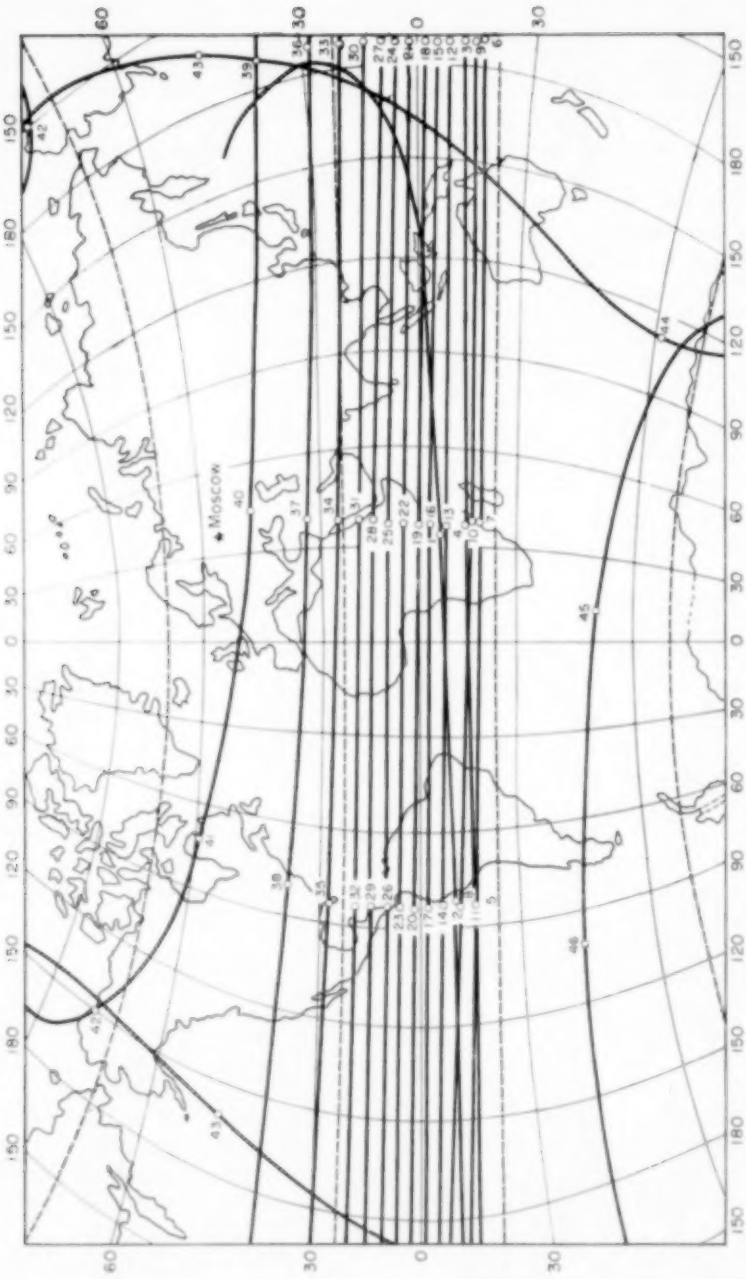


Fig. 8. Projection of the orbit of the automatic interplanetary station on the Earth.

Table 2.

Characteristic points of the automatic interplanetary station trajectory	Date	Moscow time hrs., min.	Distance from Earth's centre 10^3km	Distance from centre of Moon 10^3km
Minimum distance from Moon	6.10.1959	17·21	368·08	7·94
Photograph	7.10.1959	6.30–7.10	399·50–400·50	65·5–68·5
Greatest distance from the Earth	11.10.1959	0.44	480·50	442
Minimum distance from the Earth	18.10.1959	19.49	47·49	363
Greatest distance from Earth on the second loop	26.10.1959	20.59	489	697

two-thirds of the invisible side of the Moon being photographed. As regards photography one unfavourable circumstance was direct illumination of the surface as the result of which shadows were absent from irregularities on the surface of the Moon and the image was obtained only of the difference in reflecting properties of the different sections.

Before photographing, the orientation system was switched on, directing one end of the axis of the automatic interplanetary station towards the Sun and the opposite end with the camera objectives towards the Moon. After the automatic interplanetary station had been orientated the shutter was opened, closing the illuminator with the lunar orientation pickups and cameras. On fixing the Moon with a special photoelectric cell the photoelectric cells, reacting on the light of the Sun, were automatically switched out; immediately afterwards precise orientation was carried out of the axis of the cameras on the centre of the Moon. Photography was carried out after completing the orientation on the centre of the Moon, com-

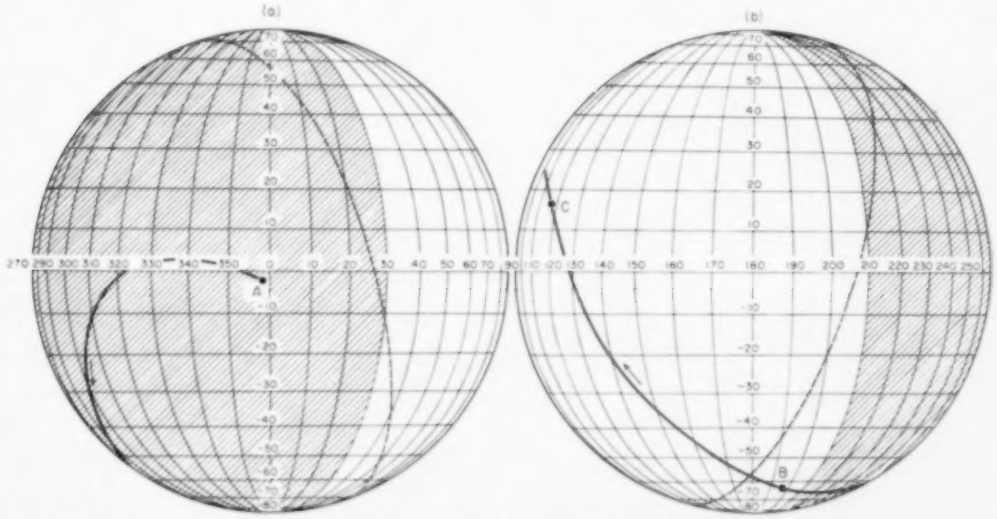


Fig. 9. Selenographic co-ordinates and projection of the orbit of the automatic interplanetary station on the surface of the Moon. *a*—hemisphere turned to the Earth; *b*—hemisphere invisible from the Earth; the shading shows the side not lit by the Sun; the continuous line is the projection of the orbit of the automatic interplanetary station; the discontinuous line is the boundary of the surface of the Moon photographed; *A*—commencement of flight; *B*—minimum distance to the Moon; *C*—point at which photography was carried out 65-68,000 km from the surface of the Moon).

mencing 7th October at 6 hrs 30 min Moscow time; at this time the distance of the automatic interplanetary station to the centre of the Moon was 65·6 thousand km. In order to obtain favourable temperature conditions inside the automatic interplanetary station, rotation of the station was applied around its transverse axis with an angular velocity equal to 2° per second after completion of the photography. Automatic development, fixation and drying of the film then followed, with transmission of the photographic image to Earth by means of a special television system.

Specimens of the photographs obtained treated up to the present are shown in Fig. 10, 11 and 12.

The photographs transmitted to Earth from the automatic interplanetary station show a small part of the Moon's surface visible from the Earth in addition to the section of the Moon invisible from the Earth. It is thus possible in places to connect the hitherto unobserved objects on the Moon's surface with those already known and to determine their co-ordinates on the Moon.

In the case of objects occurring on the edge of the visible disc the shape of which was only slightly known on account of the perspective distortions, more precise sketches have now been obtained. In other cases the objects were known only in part; the photographs obtained yielded a more precise shape for the parts situated on the visible section of the Moon together with a determination of the shape of the unknown extensions on the invisible section.

On the invisible section the presence of a mountain range was established, called the Sovietsky Mountains, together with seas and craters. One of the seas was called the Sea of Dreams in honour of the first space rocket launched on 2nd January 1959.

The photographs in Fig. 10 and 11 are the original photographic plates, from which the clearly defined interference from radio noise has been removed. The photograph in Fig. 12 was obtained from the preliminary treatment of a number of photographs and is the first attempt at composition on the basis of detected details appearing on a number of plates.

For a complete description of the invisible side of the Moon it is necessary to continue the photographic work. It is apparent that subsequent photographs with side lighting may yield new formations on the far side of the Moon.

It appears to us that the experiment of photographing in space and transmitting the pictures by radio to Earth is interesting not only from the standpoint of knowing the nature of the far side of the Moon but in addition the experiment is intensely important as a test of a new method in modern experimental. astronomy.

Fig. 11. Photograph of the Moon's surface obtained from the automatic interplanetary station.

(*To face page* 116)

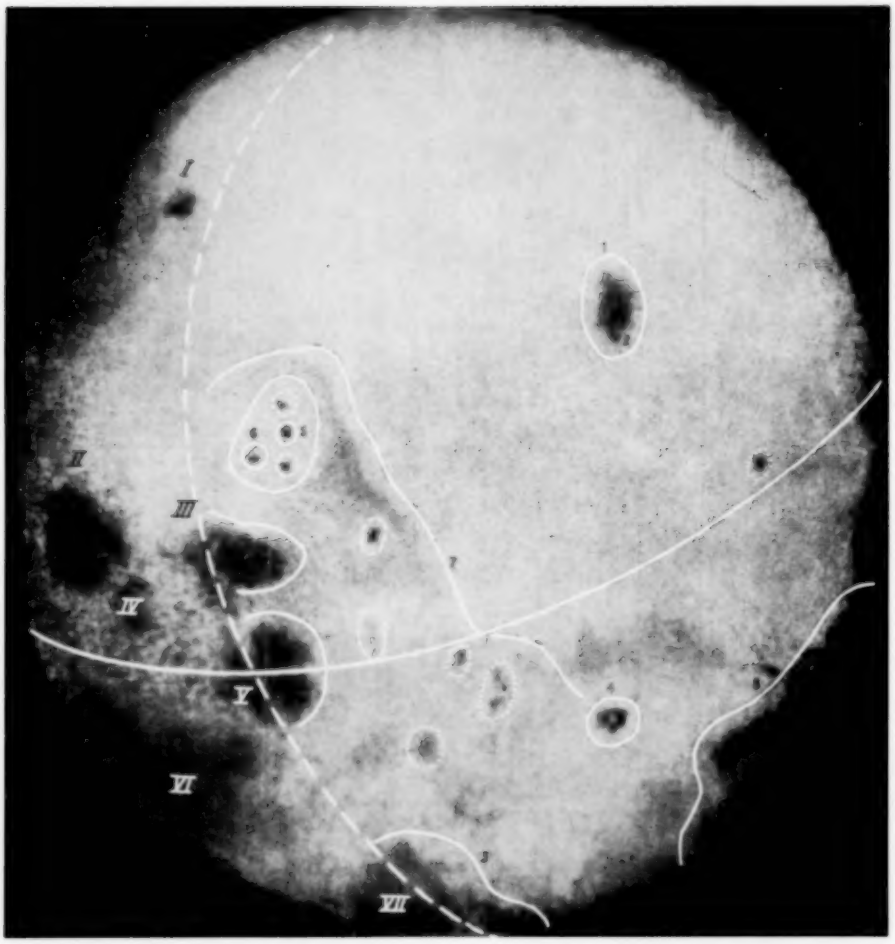


Fig. 12. Distribution on the surface of the Moon of a number of objects obtained after preliminary treatment of photographs transmitted from the automatic interplanetary station. 1—Moscow Sea (crater sea with diameter 300 km); 2—Astronauts Bay; 3—continuation on the invisible side of the Mare Australe; 4 and 5—craters from the central Tziolkovsky and Lomonosov range; 6—the Joliot-Curie craters; 7—Sovietsky Mountains; 8—Sea of Dreams; continuous line—equator of the Moon; discontinuous line—boundary of the Moon invisible from Earth; the objects outlined by a continuous line have been reliably established; the objects enclosed in discontinuous lines require precise determination of the shape; the classification of objects within the dotted lines is being established. Objects on the visible side of the Moon are shown with Roman figures; I—Mare Humboldtianum; II—Sea of Conflicts; III—Mare Marginus; IV—Mare Undarum; V—Smythii Mare; VI —Mare Foecunditatis; VII—Mare Australe.

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 117–119. Printed in Great Britain.

ABSTRACTS

No. 6 (1961)

*V. A. SARYCHEV: **The influence of the earth's oblateness on the rotational motion of an artificial satellite** (pp. 3–10).

The three right rectangular systems of co-ordinates necessary for drawing up the equations of motion comprise the absolute system of co-ordinates taking in the axis of rotation of the earth and the equatorial plane; the orbital system of co-ordinates; and a system of co-ordinates whose axes lie along the principal axes of the central ellipsoid of inertia of the artificial satellite. Expressions are derived for the projections of absolute angular speed of rotation of the orbital system of co-ordinates from Poisson's kinetic relations for the direction cosines. The equations of rotational motion of the satellite are set out in the Euler form, and simplification is applied by means of the equations of motion of the satellite mass centre in osculating elements with the usual assumption of independence of advancing and rotational motion. The equations of small oscillations of the satellite relative to the orbital system of co-ordinates for the case of arbitrary location of the orbit in space are derived by normal linearisation of the trigonometric functions.

*V. V. BELETSKII: **The classification of artificial earth satellite paths about the mass centre** (pp. 11–32).

The paper investigates the motion of the kinetic moment vector, studying the influence of gravitational perturbations and aerodynamic perturbations, taking account of the moment of aerodynamic forces as a function of the angle of attack, together with the influence of orbital regression. Twelve classes of trajectory are classified into two types—bearing respectively an aerodynamic character and a gravitational character. Aerodynamic perturbations possess a certain stability with respect to gravitational perturbations. The aerodynamic type of trajectory ceases in the presence of infinitely large gravitational perturbations as compared with the aerodynamic. At the same time the gravitational type of trajectory ceases with a finite ratio of aerodynamic perturbations to the gravitational perturbations. In the presence of interaction of gravitational and aerodynamic perturbations, in some cases the relationship of the aerodynamic moment to the angle of attack has a substantial influence. In order to investigate more closely the influence of regression of the orbital angle and the influence of regression of the perigee, aerodynamic perturbation effects are disregarded at this stage. Investigations of the rotation of the third Soviet satellite have indicated that electro-magnetic effects can also play a substantial part if power systems such as solar batteries are carried on the satellite.

*K. I. GRINGAUZ, V. A. RUDAKOV and A. V. KAPORSKII: **Apparatus for rocket measurements of the free-electron concentration in the ionosphere** (pp. 33–47).

The radio equipment employed for studying the electron concentration in the atmosphere against altitude on the vertically-launched geophysical rockets of the U.S.S.R. Academy of Sciences includes the radio-transmitter unit with transmitting aerials, the ground receiving aerials, the receiving phase-metering instruments and the recording and auxiliary equipment. The rocket transmitter generates coherent oscillations on three frequencies. The rocket aerials take the form of symmetrical linear vibrators. Reception occurred at a number of sites near to the projection on the earth of the trajectory peak. Two methods are employed for recording the measurements on 35 mm film. The first employs a loop oscillograph for recording the interference frequency and current; the second measures the phase difference, recording the characteristic currents of the Lissajous figure on the moving film. Measurement discrepancies, as indicated by the dispersion of results from variously sited receiving stations, amount to 5 per cent.

*K. I. GRINGAUZ and V. A. RUDAKOV: **Measurements of electron concentration in the ionosphere up to altitudes of 420–470 km, carried out during the International Geophysical Year by means of radio-wave transmissions from the geophysical rockets of the Academy of Sciences U.S.S.R.** (pp. 48–62).

Comparison of the sections of the curves relating electron concentration to altitude above the principal ionization maximum (over 300 km), obtained from three rocket flights, indicates that the rate of electron-concentration decay with altitude over the one geographical location may vary over a wide range (apparently according to time of year and day), thus contradicting the earlier concept of an ionization model only slightly dependent on time of day, season, latitude, etc. The variability in the fall-off rate of ionized particles with altitude in the upper ionosphere was confirmed by results obtained from ion traps installed on the third Soviet earth satellite. The results confirm a principal electron-concentration maximum in the ionosphere at altitudes around 300 km and that the so-called *E* layer of the ionosphere appears only as the result of the limited possibility of a purely altitude distribution of electron concentration on the basis of the altitude-frequency characteristics obtained from ionosphere stations. An assessment of the possibilities of using the dispersion interferometer at altitudes of some thousands of kilometers indicates their unreliability, reaching impossibility at distances greater than 20.000 km.

K. I. GRINGAUZ, V. V. BEZRUKIKH and V. D. OZEROV: **Results of measurements of positive-ion concentrations in the ionosphere employing ion traps on the third Soviet earth satellite** (pp. 63–100).

Negative ions, if they exist in the *F* region of the ionosphere, are present only in negligible quantities, since the measured electron concentrations were roughly equal to the positive-ion concentrations. Ionization of neutral particles of air as the result of the satellite motion does not appreciably affect the ion-trap measurements. The absence of appreciable quantities of negative ions in the upper ionosphere indicates that the positive ion concentrations obtained can at the same time be taken as the electron concentration values. At altitudes above 400 km there is a fairly slow diminution in positive-ion concentration, being in good agreement with the results from measurements on vertical take-off geophysical rockets up to 470 km. At 1000 km a positive-ion concentration of approximately $5{\cdot}7 \times 10^4 cm^{-3}$ of ions of a given mass is appreciably greater than the concentration of all ions with other masses, confirming mass spectrum measurements according to which atomic oxygen is the main constituent of the ionosphere at this altitude.

*K. I. GRINGAUZ, V. V. BEZRUKIKH, V. D. OZEROV and R. E. RYBCHINSKII: **The study of interplanetary ionized gas, high-energy electrons and corpuscular radiation of the sun, employing 3-electrode charged particle traps on the second Soviet space rocket** (pp. 101–107).

At distances up to four earth's radii from the earth's surface, plasma is observed with a temperature not greater than tens of thousands of degrees. The existence of plasma at such distances confirms results obtained from the first space rocket in January 1959 and on the third space rocket in October 1959 (up to 7000 km). In the range 55,000 to 75,000 km, electron fluxes of approximately $10^8 cm^{-2} sec^{-1}$ were obtained with energies in excess of 200 eV, confirming the results from the first space rocket. From 9.30 a.m. Moscow time, 13 September 1959 up to impact with the moon, positive ions were recorded with energies in excess of 15eV at a concentration around $2 \times 10^6 cm^{-2} sec^{-1}$. At various distances from the earth, for example at 125,000 km, a proton flux was observed with energies in excess of 25 eV in the course of telemetering data during the flight of the automatic interplanetary station in October 1959.

*K. I. GRINGAUZ, V. G. KURT, V. I. MOROZ and I. S. SHKLOVSKII: **Ionized gas and fast electrons in the vicinity of the earth and in interplanetary space** (pp. 108–112).

The experimental data from the second Soviet space rocket has been analysed in four sections, comprising the first section up to 22,000 km from the earth's surface, from 22,000 km to 50,000 km, thence up to 70,000 km and greater than 70,000 km. For assessing the ion concentration from collector-current measurements importance attaches to the electrical potential of the container relative to the medium, depending on several factors, including the fast-electron flux in the second radiation belt, the maximum radiation of which occurs at a distance corresponding to the first of the four sections. Hence special attention is paid to the current variations in the "+15V" trap. A study of the relationship of current to electron concentration for different plasma densities leads to the conclusion that the electron fluxes corresponding to energies greater than 200 eV in the region of the upper radiation belt do not exceed $2 \times 10^7 cm^{-2} sec^{-1}$, thus contradicting the concept of high electron fluxes with energies 20–30 keV in the maximum of the upper radiation belt. An assessment of plasma ion concentration in conjunction with the ion concentration as a function of distance from the earth's surface indicates that the plasma in question is not interplanetary ionized gas, but forms an extensive cloud, being an ionized component of the outermost section of the earth's atmosphere, termed geocorona, displaying a marked increase in the geocorona plasma density starting from 15,000 km, whilst at lesser distances the density variation is negligible. It is proposed that the polarised component of zodiacal light is governed by diffraction of sunlight on dust particles, and not on free electrons as proposed by Behr and Siedentopf.

*V. I. KRASOVSKII, S. SHKLOVSKII, YU. I. GAL'PERIN, E. M. SVETLITSKII, YU. M. KUSHNIR and G. A. BORDOVSKII: **The detection of electrons with energies of approximately 10 keV in the upper atmosphere** (pp. 113–126).

The study of aurora spectra has pointed to their origin from protons with high velocity dispersion. Frequently hydrogen emission was lacking. In this case an explanation was proffered on the basis of the presence of not particularly hard electrons with energies approaching hundreds and thousands of electron volts. An attempt to detect these not-particularly hard electrons on the third artificial satellite for the first time yielded information regarding the existence at altitudes up to 1900 km of high electron concentrations with energies predominantly around 10 keV. Electrons were recorded travelling earthwards and at small angles to the lines of magnetic force, appearing at altitudes above 1900 km. It was established that the electron energy flux with magnetic moment such as to allow penetration into the *F*-layer might attain values of approximately 1 erg $cm^2 sec^{-1}$. A study of basic data reflects the higher sensitivity of the apparatus to high-energy electrons, so that for non-monochromatic electrons it is to be expected that the maximum electron flux recorded would correspond to energies below 14 keV. The relationship

of equivalent electron energy to the angle between the indicator axis and the lines of magnetic force shows only a slight dependence on the direction of the electron path. The lowest equivalent energies are observed in the polar regions. Variations in time may sometimes be appreciable over intervals markedly less than the circling time of the satellite.

*V. G. ISTOMIN: **The variation in positive-ion concentration with altitude from data of mass-spectrometry measurements on the third satellite** (pp. 127–131)

Up to altitudes of 550 km the variation in positive-ion concentration is in satisfactory agreement with the variation in electron concentration, whilst above 650 km mass-spectrometry measurements show a wide diversity on three different days. A comparison of mean concentration measurements obtained by mass spectrometry with experimental data employing ion traps provides confirmatinon of the $O+$ concentration as a function of altitude, and of electron concentration as a function of altitude. The agreement between the N^+, O^+, N_2^+, NO^+ and O_2^+ positive-ion concentration against altitude and the electron concentration curve points to the fact that light ions, with mass numbers beyond the range of the instrument (hydrogen ions H^+, H_2^+, and helium He^+) do not exist in appreciable quantities and the ionosphere up to 900 km remains oxygen-nitrogen. With the presence of light ions it is natural to expect an increase in their relative concentrations with altitude, as is found for ions of atomic nitrogen in relation to ions of atomic oxygen and for ions of atomic oxygen in relation to still heavier ions/1,2/. An increase in relative concentrations of ions H^+, H_2^+ and He^+ would thus lead to a divergence between the electron concentration and the ion concentration.

*L. V. KURNOSOVA, L. A. RAZORENOV and M. I. FRADKIN: **Short-term increases in the intensity of the nuclear component of cosmic rays associated with solar activity** (pp. 132–138).

Equipment installed on the second space rocket comprised two independent instruments, each in the form of a Cerenkov counter designed for recording nuclei in cosmic radiation, travelling at relativistic speeds. More than 100 nuclei with $Z \geqslant 15$ were recorded, over 3000 with $Z \geqslant 5$ and approximately 30,000 nuclei with $Z \geqslant 2$. Changes in the flux intensity of different groups of nuclei of cosmic rays were plotted and checked against solar activity on 12 September 1959. Whilst it is difficult as yet to point to the concrete mechanism of heavy-nuclei acceleration on the sun, the manner in which the nuclei are ejected from the sun's surface, and to the connection between these processes and radio-emission, the observed phenomena are extremely unlikely to be fortuitous coincidence and point rather to an extremely probable supposition regarding their interrelationship.

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 120–121. Printed in Great Britain.

RESEARCH NOTES

The soft X-ray absorption of aluminium and rocket measurements of solar radiation

(*Received* 6 *June* 1961)

A programme of measurements of soft (2-20 Å) X-ray absorption coefficients has been instigated in connection with the development of gas photon counters for use in the U.K. Space Research Programme. The need for revision of generally accepted values has been shown and the first results of this investigation are being published elsewhere[1]. The purpose of the present note is to point out the effect of the revised absorption coefficients of aluminium on previous rocket measurements of solar X-ray emission with aluminium window photon counters.

The new data has been obtained with a curved crystal vacuum monochromator and a proportional counter detector and care has been taken to avoid the small angle scattering effects pointed out recently by Parratt[2]. It is considered that this experimental method is capable of greater precision of absorption measurement than those used previously[3, 4, 5].

Aluminium is of particular importance since it remains the one material suitable for photon counter windows to operate in the wavelength range 10-20 Å. The *L*-shell absorption coefficients of aluminium, above the *K*-edge at 7·9 Å, have been found to lie consistently nearly 25 per cent higher than Allen's[4] generally accepted values for this metal. The photon detection efficiency of the gas counters used in previous rocket measurement of the solar X-ray spectrum is given[6] for any wavelength λ by the product

$$(\text{P.D.E.})_\lambda = W_\lambda \times G_\lambda$$

where W_λ is the fractional transmission of the window material and G_λ the fractional absorption of the gas filling to incident radiation of wavelength λ. Thus a revision of W_λ has a direct effect on the solar flux computed from any measured count rate.

To indicate the order of the correction necessary for most of the previous rocket measurements, a calculation has been carried out for the data of flight Aerobee 14 on 1st December 1953, when the Sun was "unusually quiet." This result provides the minimum solar X-ray flux ever measured for the wavelength region below 20 Å. A photon counter fitted with an aluminium window of 1·59 mg cm^{-2}, with effective sensitivity in the wavelength band from 8-20 Å was flown in this rocket and from the measured count rate of this detector above the absorbing layers of air, with the assumption of a solar spectral emission as of a grey-body radiator at $2 \cdot 10^{6\circ}$ K, the N.R.L. group determine[6, 7] a solar flux in the 8-20 Å band of $3 \cdot 9 \times 10^{-4}$ erg cm^{-2}/sec. Here the absorption coefficients of Allen[4] were used. Re-calculation of this result using the new absorption data gives a value for this "minimum" solar X-ray emission of $5 \cdot 3 \times 10^{-4}$ erg cm^{-2}/sec.

Other rocket measurements of this part of the solar spectrum are altered by the same factor, provided that the grey-body spectrum of $2 \cdot 10^{6\circ}$ K is retained. With the assumption of a lower radiation temperature (as seems likely[8, 9] with the "very quiet" Sun) the adjustment is increased.

K. A. POUNDS

Department of Physics,
University of Leicester,
England

REFERENCES

1. B. A. COOKE, K. A. POUNDS, P. C. RUSSELL and E. A. STEWARDSON, *Proc. Phys. Soc.* (*to be published*).
2. L. G. PARRATT, J. O. PORTEUS, H. W. SCHNOPPER and T. WATANABE, *Rev. Sci. Instrum.* **30**, 344 (1959).
3. E. JÖNNSON, Thesis, Uppsala (1928).
4. S. J. M. ALLEN, *X-rays in theory and experiment* (ed. A. H. Compton and S. K. Allison), Van Nostrand, New York (1935).
5. H. H. BIERMANN, *Ann. Phys. Lpz.* **26**, 740 (1936).
6. E. T. BYRAM, T. A. CHUBB and H. FRIEDMAN, *J. Geophys. Res.* **61**, 251 (1956).
7. H. FRIEDMAN, *The Upper Atmosphere* (ed. J. A. Ratcliffe), Academic Press, New York (1960).
8. M. WALDMEIER, *Z. Astrophys.* **30**, 137 (1952).
9. H. C. VAN DER HULST, *Bull. Astr. Insts. Netherlds.* **11**, 135, 150 (1950).

Linear arrays of small craters

Received 23 *October* 1961.

In discussions of geological processes on the Moon, frequent reference is made to various linear arrays of small craters. Such arrays occur, for example, in the vicinity of Copernicus, as well as at several locations in the highlands, such as the edge of the crater Müller. It is often pointed out that such linear formations are difficult to explain as resulting from random impacts. On the basis of this difficulty, the linear arrays are commonly held up as evidence of geologic activity on the Moon[1, 2]. That is, some subsurface phenomenon gave rise to a nearly linear crack on the surface, along which volcanic activity occurred, which resulted in craters also arrayed along a line.

It appears that an alternate explanation is available to account for such linear arrays—an explanation which does not require the hypothesis of volcanic activity or tectonic forces capable of producing such cracks. It is possible that the dynamic behaviour of meteorites on an approach trajectory to the Moon might give rise to linear arrays of impact craters.

Like every other body in the solar system, the Moon has its own Roche limit such that when bodies approaching the Moon come within this limit they can no longer hold themselves together by reason of their own gravitational forces. Thus, during a portion of the trajectory of such a body inside a Roche limit of the Moon, the body would disintegrate into pieces whose cohesive force is chemical rather than gravitational; that is, monoliths. Throughout the remainder of the flight, these monoliths would be dispersed from each other, on the basis of such dynamical parameters as the initial spin rate of the original body, the trajectory in the Moon's field, and the perturbing forces introduced by the Earth and gravitational anomalies on the Moon. In addition, bodies which pass close to the Earth and then proceed to strike the Moon would be separated by their passage through the gravitational field of the Earth. Then when finally striking the Moon, this separation distance would be considerably greater than if the separation had been caused only by their entry into the Moon's gravitational field.

Such a disintegration must occur for any large body approaching or passing the Moon or any other planet. The problem is: How far apart would the separate pieces get before impacting the surface? It would appear that this question could be resolved by calculations which would show at least the range of parameters which would be required in order to obtain separations such as those observed in the linear arrays of craters. This explanation might also account for some of the multiple central peaks observed in some of the large craters.

A. R. Hibbs

Division of the Space Sciences,
Jet Propulsion Laboratory,
California Institute of Technology,
Pasadena

REFERENCES

1. R. B. Baldwin, *The face of the Moon*, University of Chicago Press, Chicago (1949)
2. H. C. Urey, *Endeavour*, Vol. 29, No. 74, 87 (1960)

Planet. Space Sci. Pergamon Press, 1961. Vol. 8, pp. 122–123. Printed in Great Britain.

BOOK REVIEWS

RICHARD N. THOMAS and R. GRANT ATHAY: **Physics of the solar chromosphere.** Interscience Publishers, New York, London. 1961. Pp. x+422. $15.50.

This important monograph is not a general review of the problems connected with the solar chromosphere, but is a highly specialised work dealing with a particular approach to these problems. The fundamental concept suggested by the authors and their co-workers is that the chromosphere is not in local thermo-dynamic equilibrium (L.T.E.).

The first four chapters lay the foundation for the remainder of the book. Individual chapters deal with a general review of the problems of the chromosphere, the observational data, the structure of the chromosphere and the mathematical theory of non-L.T.E. effects. The model proposed for the chromosphere is discussed in Chapter 5. The subsequent three chapters deal with the application of the theory to the chromosphere; first, to determine the scale height distribution of electron density and temperature and then to include helium emission and radio noise data. Chapter 9 considers data which is affected by both self-absorption and non-L.T.E. effects and the book is concluded by a summary chapter. Two appendices give the Balmer continuum and line data for a normal chromospheric region and short notes on the preceding chapters. There is an extensive bibliography and a good index. Each chapter of the book has been the responsibility of one or other of the authors—the particular author being given with the chapter headings.

This book, which has been concisely and excellently written, should become a standard reference work on the treatment of the solar chromosphere from a non-L.T.E. point of view. Since the fundamental theory is so well presented this book should have considerable value for other non-L.T.E. studies in astrophysics and in the field of thermonuclear discharges.

D. MCNALLY

T. E. STERNE: **An introduction to celestial mechanics.** Interscience Publishers Inc., New York and London. 1960. Pp. xi+206. Cloth 34/-; Paper 19/-.

This book is based partly on the dynamical section of a half-year astronomical course for graduate and undergraduate students at Harvard, and partly on the author's own research. The first of the six chapters is devoted to the two-body problem and is followed by a set of examples, one of which briefly introduces the concept of a sphere-of-influence. The second chapter contains a discussion of the external potential of an irregular body, but the data on the earth is unfortunately obsolete. Chapter 3 discusses the various coordinate systems used by astronomers, while Chapter 4 discusses the general perturbations of an orbit. Most of Chapter 5, which investigates perturbations due to the earth's oblateness and due to air resistance, is based on the author's own work. The final Chapter is devoted to the numerical techniques used in orbital computations.

The book is more lucid than many texts on celestial mechanics and should provide a useful introduction to the subject, both for students and for those concerned with the practical problems of space research. Owing to the nature of the topics discussed, however, the book will have a greater appeal for workers on near-satellites than for those associated with other types of space vehicle. It is to be regretted that the author, being preoccupied with his own results, makes little reference to other people's theories, which are sometimes more comprehensive.

G. E. COOK

I. S. SHKLOVSKII: **Cosmic radio waves.** Translated by Richard B. Rodman and Carlos M. Varsavsky. Harvard University Press, Cambridge, Mass., 1960. Pp. 444. $12.50.

Dr. Shklovskii is the Director of the Department of Radio Astronomy of the Sternberg Institute of Moscow University. He is well known as an original and enthusiastic astrophysicist and has made many valuable contributions to radio-astronomy. It is much to be regretted that several years have elapsed since the publication of this excellent book in Russia and its appearance in an English translation. The translation is good and a determined effort has been made to bring the book up to date by revision in 1958 and by an appendix added in 1960.

This is the first book on radio-astronomy which does not attempt to cover the whole subject and pay the inevitable price of superficiality. It is devoted solely to the radio emission from objects beyond the solar system and excludes, for example, the subject of the Sun. It is written without much mathematics and is an easy book to understand.

The first third of the book includes a brief chapter on techniques which is rather perfunctory, and is followed by a detailed account of the observational data. This second chapter is profusely illustrated and provides a reference work of considerable value. Most of the observations described were made in Australia, Great Britain or the U.S.A.; at the time of writing the volume of observational work from Russia was small.

The remaining two-thirds of the book are concerned with radio emission from the Galaxy, the hydrogen-line, the role played by the cosmic rays, and extra-galactic radio emission. Dr. Shklovskii has for many years set the pace in interpreting the observations of radio-astronomers. As one would expect, a sub-

stantial part of these chapters is original and to a large extent controversial. Nevertheless, although some of the detailed conclusions must be modified by more recent data, they give a valuable and interesting account of current ideas on the part played by supernovae and the interstellar medium in the generation of cosmic radio waves and their relation to cosmic rays. The discussion relates these ideas to the distribution of radio emission from the Galaxy, and also to the generation of radio emission in radio galaxies. In particular many physicists should find the discussion of cosmic rays and the interstellar medium well worth reading.

One particularly valuable task which Dr. Shklovskii has achieved, a task which could hardly have been performed better by a visitor from outer space, is to discuss Russian and Western researches impartially and to make them into one story: on these grounds alone this book should be in every scientific library. It should also be read, not only for its scientific value, but because it is written with a lively enthusiasm and a sense of the excitement of scientific discovery.

R. HANBURY BROWN

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 125–141. Printed in Great Britain.

SOME RESULTS OF INVESTIGATIONS OF AURORAE AND NIGHT AIRGLOW DURING THE I.G.Y. AND I.G.C.

V. I. KRASOVSKII
The Institute of the Physics of the Atmosphere, the U.S.S.R. Academy of Sciences, Moscow

(*Received* 24 *July* 1961)

Abstract—A general review is given of the results of the airglow and auroral investigations during I.G.Y and I.G.C. in the Institute for Atmospheric Physics of the Academy of Sciences of the U.S.S.R. A strong helium emission at 10830 Å has been discovered in sunlit aurorae and in the ordinary twilight airglow. It is now observed regularly. A detailed theory of this fluorescent emission has been developed. Extensive observations on the hydroxyl bands between 5000 and 12,000 Å have been carried out at a number of stations. Marked seasonal and geographical variations have been found in them. During intense aurorae the line-width of the 6300 Å emission from great altitudes increases. The effect has been studied using Fabry-Perot interferometers. Contributions have been made to knowledge on auroral morphology. The interpretation of some of the results which are described in the review are discussed.

The main objective of the Institute for Atmospheric Physics of the Academy of Sciences of the U.S.S.R. during the I.G.Y. and I.G.C. (the International Geophysical Year and International Geophyscial Co-operation) was the investigation of elementary processes in the upper atmosphere. As a result of observations of aurorae, the twilight flash and the night airglow, rich material has been obtained about the nature of these processes. Observations were made at Loparskaya ($\Phi = 63°37'$, $\Lambda = 126°42'$), Roshchino ($\Phi = 56°35'$, $\Lambda = 116°46'$) and Zvenigorod ($\Phi = 51°03'$, $\Lambda = 120°16'$)[1]. However, numerous details of our investigations have not yet been carefully compared and examined. The first attempt in this direction is being made in this paper. Somewhat arbitrarily the material is divided into two sections: "Aurorae" and "The Night Airglow." Since observations were made during a period of maximum solar activity, the conclusions drawn should be most characteristic of this period. More comprehensive conclusions will become possible when additional investigations are made during years of minimum solar activity.

AURORAE

The term "aurora" has commonly denoted the appearance of some luminous formations distinguishable by the naked eye, mainly with definite configurations. However, by using electrophotometry over the celestial sphere N.V. Jorjio[2] showed that only a portion of the emission is concentrated in definitely shaped formations. In the majority of the cases the integrated emission of the surrounding background, mainly homogeneous in intensity, was found to be greater. This emission was not observed visually due to the low contrast sensitivity of the eye when the illumination is poor. The maximum value of the recorded ratio of total luminous energy flux of the background to that of definitely shaped forms was 10 and the minimum value was 0·05 (which is characteristic only of very bright formations of short duration). The polar part of the celestial sphere at Loparskaya even on quiet nights radiated fairly regularly a N_2^+ emission detected by spectographs. During the I.G.Y. and I.G.C. N. V. Jorjio on a few occasions succeeded in observing the glow in the north not seen by the eye. However, after electrophotometry had been made along the meridian it was shown that at altitudes of 100-150 km, beginning with the geomagnetic latitude of the centre of auroral zone, 67°, a steady increase of the night airglow luminosity was recorded. The airglow intensity near the northern horizon was approximately one and a half times that of the zenith at Loparskaya. If the excess intensity could be

concentrated in the zenith in an arc of an angular width of 1°, then the arc would be of approximately the second category. In the course of the I.G.Y. and I.G.C., 35 aurorae were observed in Zvenigorod during magnetic activity (with K-index > 4). They were photographed by wide-angle cameras only 14 times due to the limited sensitivity of these instruments. However, the spectrograph recorded an increased intensity of the oxygen red line 6300 Å on 157 occasions. Because of bad meteorological conditions and bright nights there was no possibility of making observations on 25 magnetically active nights with K-index > 4. It should also be mentioned that Barbier, Roach and others[3, 4] have found that in equatorial regions, red arcs invisible to the eye frequently occur near the zenith. They emit 6300 Å of atomic oxygen. Apparently aurorae of moderate activity are characteristic of any geomagnetic latitude.

Thus, a radical revision of the term "absence of aurora" in high and low altitudes is necessary. In view of the above considerations any statistics of aurorae seem to be reasonable only for aurorae of equal intensity and colour. To characterize the auroral type it is not enough to indicate the form which is most striking to the eye; it is also necessary to note its relative contribution to the total auroral emission, the degree of development of the brightest, most definitely and finely shaped formations, e.g. elementary small rays. In an aurora ionization of the atmosphere takes place in addition to excitation. For the macroscopic properties of the upper atmosphere, such as the general volume ionization which determines the propagation and absorption of radio waves and the surface electric conductivity essential for the creation of geomagnetic disturbances, the most characteristic feature is a correlation not only with separate isolated formations along magnetic field lines but with vast diffusion fields. Intense geomagnetic and ionospheric anomalies usually correlate with the appearance of ray forms which themselves cannot be accompanied by fields of full ionization and surface conductivity. The true cause of this dependence lies in the accompanying diffuse auroral background.

Frequent occurrence of usually unobserved diffuse aurorae causes averaging of various types of spectra several different formations being found on the line of sight. The averaging is increased by atmospheric scattering of light, especially in the presence of even slight cloudiness of the cirrus type. Therefore for conclusions to be reached it is necessary to use only data obtained in very clear weather. But even in favourable conditions the averaging occurs because of rapid changes in the type of aurorae and their angular motions. This defect is especially evident in photographs of spectra obtained with long exposures. It is less when fast electrophotometers and spectrometers are used, but distortions may appear in the spectral distribution of an emission recorded due to rapid variations of auroral intensities. The difficulties indicated above prevent us from obtaining definite laws governing observed intensities in auroral spectra. However, even a discovery of a trend may lead to important conclusions.

Auroral hydrogen emission was usually regarded as a result of an indirect interaction of primary protons of solar origin with upper atmospheric atoms and molecules. It was believed that such protons appear only at the time of great geomagnetic disturbances. But, as has already reported by V. S. Prokudina[5] a regular registration of a narrow contour hydrogen emission line from the night sky was made for the first time. This was achieved by means of high-quality spectrographs with a dispersion up to 80 Å mm^{-1} and a resolution up to 1 Å. Further details of this emission will be reported in the section entitled "The Night Airglow," but here we shall focus our attention on the fact that by means of similar high-quality spectrographs G. I. Galperin, N. N. Shefov and F. K. Shuiskaya[6–11] and others have recorded rather frequent appearance of H alpha and H beta hydrogen emission with a wide contour which in the magnetic zenith has the Doppler shift to the blue side of the spectrum. Our Institute has a unique collection of spectra with such hydrogen emission. Throughout the section on

"Aurorae," wide contour hydrogen emission will always be meant, but for the sake of brevity this will not be specifically stated.

As a result of G. I. Galperin's observations at Loparskaya it was established that in the hydrogen emission from the magnetic zenith the Doppler shift to the short-wave side of the spectrum amounts to about 300-400 km/sec in the region of the contour maximum intensity. The shift may be traced to 1500-2000 km/sec. In these contours some Doppler shift to the red side of the spectrum is also observed. This is probably due to hydrogen emission at high magnetic zenith angles near the horizon and scattered by lower atmospheric levels. The Doppler shift increases under conditions of poor visibility. It is remarkable that within the error ± 100/sec the main features of the contour remain constant in all aurorae. F. K. Shuiskaya established that the similar picture is characteristic of hydrogen emission at Roshchino. At Zvenigorod only one very weak appearance of the hydrogen emission in the magnetic zenith was recorded. Therefore it is possible to judge only about the contours of this emission in the magnetic horizon. But even in this case data obtained at Zvenigorod, Roshchino and Loparskaya coincide.

The above observational data make it possible to state that the emission recorded is due to hydrogen atoms with low non-monochromatic velocities. Their velocities differ essentially from higher monochromatic velocities of solar corpuscular streams about 10^3 km/sec. Slow radiating hydrogen atoms could appear as a result of the retardation without emission of primary fast protons from solar corpuscular streams or as a result of the acceleration of atmospheric protons during some processes of interaction of these streams with the Earth's outer atmosphere. G. I. Galperin[6-9] has made a systematic comparison of hydrogen emission intensities at the magnetic zenith and on a horizontal white screen. It transpired that these intensities are practically the same even if there is some bright auroral formation at the zenith. From these observations the conclusion is drawn that the hydrogen emission region is diffusive, that it occupies a considerable portion of the celestial sphere and apparently is not associated or is very weakly associated with the brightest and most sharply outlined patterns of auroral displays. Hydrogen emission regions as well as bright formations regularly move along the magnetic meridian to the south in the evening and return to the north during morning hours. On the average at midnight the hydrogen emission at Loparskaya was observed more frequently to the south of Loparskaya, though sometimes it appeared in the north.

Earlier Meinel[12] has described the unique case of near coincidence of zenith angles of hydrogen emission intensity maximum with the intensity maximum of other emissions in a quiet arc observed near the horizon. G. I. Galperin was eager to find similar data. Even by using exposures down to 4 min he obtained only doubtful results. In consequence O. L. Vaisberg[13,14] repeated at Loparskaya investigations on the hydrogen emission using a fast spectrometer. With a slit providing a resolution of about 10 Å, the spectral region of 2000 Å was scanned for 30 sec with the threshold sensitivity to the hydrogen Hβ emission in 20 rayleighs. It was found that the hydrogen emission never increases in any bright forms near the zenith. An increase is observed only in the formations near the horizon, especially the southern horizon. However, this can be explained in reality by the superimposition of more distant high fields of the hydrogen emission on closer, lower fields of other emissions. The increase of the intensity of the diffuse hydrogen regions near the horizon is due to the Van Rhijn effect; the decrease just at the horizon is due to atmospheric absorption.

Thus, at present no increase of the hydrogen emission in zenith forms was found and there are no convincing proofs of an increase in any formations near the horizon. The assumption of the reality of the coincidence near the horizon of the intensity maxima of hydrogen and other emissions with the exclusion of it near the zenith is absurd since it would connect the place of coincidence with the position of an observer. At present it seems more probable that the hydrogen emission in general does not

increase in such sharply defined auroral formations as arcs, bands and rays. It should be recalled that the diffusive character of the hydrogen emission regions was considered inevitable earlier, in the light of the charge exchange mechanism first suggested by I. S. Shklovskii[15]. According to this mechanism, now generally accepted, primary protons take part in charge exchange with neutral atoms and molecules of the Earth's atmosphere. However, having become neutral hydrogen atoms they stop spiralling along geometric lines of force. The very interesting suggestion has been made by I. S. Shklovskii[16] that some protons become neutral atoms in interplanetary space in their path to the Earth.

According to G. I. Galperin, at Loparskaya the hydrogen emission appeared either before the beginning of aurorae or during any stage of their development. Sometimes it was observed during several successive days but it was not recorded of any preferable time of day. It is worth noting that on a number of occasions at Loparskaya an intensive hydrogen emission at a magnetically quiet time was observed without appearance of any visually-observable auroral forms. However, in the majority of cases it began several hours before their appearance. At Roshchino and Zvenigorod the hydrogen emission appeared only occasionally during some very strong geomagnetic disturbances. However, at Roshchino as well as at Zvenigorod it was not recorded during many other very strong geomagnetic disturbances. Yet at these stations, provided favourable meteorological conditions existed, an aurora always appeared, detected visually or by means of spectrographs, during geomagnetic disturbances and there was no hydrogen emission in the spectrum. Data obtained by G. I. Galperin and O. L. Vaisberg indicate that hydrogen emission regions have longitudinal extension of the order of hundreds of kilometres and sometimes of more than a thousand kilometres with maximum intensity of thousands of rayleighs. It is evident that the total energy corresponding to such extended emission regions is great. From the north these regions are surrounded by sharply outlined, presumably movable, auroral formations without hydrogen emission. In their turn observations at southern stations indicate unambiguously that they are surrounded by similar formations from the southern side too. For instance, at Roshchino an aurora with an ionized nitrogen molecule emission was recorded with spectrographs at the magnetic zenith 39 times, but simultaneously the hydrogen emission in the magnetic horizon was observed only 10 times and at the magnetic zenith only 3 times. More to the south, in Zvenigorod, an aurora with an ionized nitrogen molecule emission was recorded at the magnetic zenith 37 times, but simultaneously the hydrogen emission in the magnetic horizon was observed only 7 times and at the magnetic zenith only once and with a very low intensity.

The hydrogen emission region along the parallel was found to be more or less uniform. It usually increases somewhat towards the horizon due to Van Rhijn effect. Though observations of the hydrogen emission by means of the spectrometer have considerably lower angular resolution than have visual and photographic detections of intensity irregularities of bright auroral forms along the parallel, nevertheless, one cannot ignore the absence of intensity fluctuations exceeding 10-20 per cent of the average above the level which is due to Van Rhijn effect.

Data obtained by G. I. Galperin by means of spectrographs show that the hydrogen emission most frequently exists in spectra in which atomic emissions predominate and that it is relatively seldom observed in spectra with strongly developed molecular emissions. The possibility is not excluded that in a number of cases the hydrogen emission is lost only due to strong blending and not as a result of its actual absence from the latter. O. L. Vaisberg pointed out that during the appearance of the hydrogen emission in faint aurorae there is apparently a linear dependence between all emissions. As is clear from the following, this peculiarity is characteristic of high-altitude aurorae and indicates the predominant appearance of the hydrogen emission at great heights. However, the hydrogen emission is not an in-

separable component of atomic spectra. It is completely absent in many of them.

Analyzing valuable material obtained by means of the spectrometer O. L. Vaisberg established that in fields of hydrogen emission the ratio of the intensity of the N_2^+ Meinel system to the intensity of the N_2 first positive system is almost by one and a half times lower than in the absence of hydrogen emission. On the other hand N II emission near 5004 Å during similar comparisons is one and a half times greater in the majority of cases. The difference between auroral spectra with the hydrogen emission and spectra without such emission ends here. There is nothing surprising in this since a similar situation is observed in laboratory spectra[67]. At present it is difficult to evaluate the extent of blending of emissions caused by protons with emissions caused by other agents. However, the lack of a definite dependence of the hydrogen emission on other emissions undoubtedly indicates the superimposition of emissions due to various mechanisms.

Analyzing numerous data G. I. Galperin, with spectographs, and O. L. Vaisberg, with the spectrometer, have established that the spectra of diffusive aurorae without hydrogen emission do not differ in the visible region from the spectra of other sharply outlined forms. It should be noted that their conclusion refers to aurorae in parts of the upper atmosphere unlit by the Sun. However, aurorae with anomalously increased oxygen red emission at 6300 Å are excluded. In such aurorae this emission may be by an order or several orders of magnitude more intense than, for instance, the 5577 Å oxygen green emission. Usually the rays, spots and surfaces of these aurorae are tipped with bright red, the colour being apparent to the naked eye. Let us provisionally call this type the zeroth. All other spectra, irrespective of the kind of luminous formations, can be provisionally divided into three types between which, of course, there are no sharp graduations. To the first type belong those spectra of the night airglow in which a manifold increase of the oxygen red emission at 4300 Å and the correlating nitrogen green emission at 5200 Å is observed. Simultaneously weak bands of the ionized nitrogen molecule appear in the ultraviolet. The corresponding type of aurorae is most frequently observed in low latitudes. In the next type of spectra many emissions of neutral and ionized atoms are observed and, in addition, traces of ionized and neutral nitrogen molecule emission appear in the visible region. Aurorae with such spectra are often observed not only in low latitudes, but also in high latitudes. Lastly, the third type is characterized by numerous intense molecular bands. It is remarkable that in such spectra the green line of OI and the first and second positive and negative systems of N_2 mutually correlate. Aurorae which correspond to them most frequently appear in high latitudes. The difference between the spectra described are apparently accounted for by the depth of penetration of the exciting agent into the atmosphere. The first type of spectra corresponds to the highest atmospheric regions with mostly dissociated oxygen and nitrogen; the second type to the regions where molecular nitrogen is more abundant and the third to lower regions where nitrogen is the principal constituent. Traces of bands of ionized molecular oxygen appear in spectra from very low and intense forms. However, to avoid misunderstanding, some remarks about the low red edge of aurorae must be made. Traces of an O_2^+ emission are very weak and cannot explain the undoubted visible coloration of the base of such aurorae. Since the average lifetime of the initial state of atomic oxygen for emission of the green line (about 1 sec) is much greater than the average lifetime of the initial state of molecular nitrogen for emissions of the first positive system (about 10^{-6} sec), the 5577 Å emission at an altitude lower than 100 km will be inevitably quenched by frequent collisions, as, in higher regions, is the 6300 Å emission, with the average lifetime of the initial state of 100 seconds. This was confirmed by spectra of auroral red lower borders obtained with exposures of 10 to 30 seconds by G. I. Galperin.

In auroral spectra of the molecular type there are always atomic emissions though their rela-

tive intensity is not high. Sometimes intense luminous formations with spectra of the zeroth, first and second types, appear over aurorae with spectra of the third type. They may last for several minutes or more, and may exist simultaneously. Such superimpositions are most easily seen in ray forms. They testify to the fact that the trajectory of the exciting agent penetration is the same for the low and for the upper parts of ray structures and that the intense superimposition observed in the auroral structure is created only by the agent which is characterized by less deep penetration. Aurorae not accompanied by the hydrogen emission are undoubtedly caused by electrons. The lower boundary of aurorae indicates the thickness of the atmosphere which can be traversed by these electrons. Thus, the minimum energy of electrons, which are characterized by the deepest penetration, can be evaluated in normal aurorae as 10 keV or a little bit lower. It should be remembered, however, that the depth of penetration of electrons with higher energies can depend on the angle of their motion relative to the geomagnetic line of force. In this case the superimposition in the auroral structure could be attributed to the anomaly in the distribution of electron motion directions relative to the geomagnetic line of zone.

V. S. Prokudina[18, 19], N. V. Jorjio[20] and especially O. L. Vaisberg have noted an interesting dependence between some permitted nitrogen and oxygen atomic emissions as well as between not very intense oxygen red emission 6300 Å, on the one hand, and emissions of the first and second positive and negative systems of molecular nitrogen, on the other. In spectra of the first type a linear correlation is revealed between all the above indicated emissions. Spectra of the third type are characterized by a relative decrease of the intensity of the first group of emissions relative to the second group together with an increase of the total auroral intensity. All emissions of the second group linearly correlate with each other. Though the scatter of true values compared with the average is within one order of magnitude, nevertheless, the average ratios of the intensity of the emission at 6300 Å to that of the emission at 5577 Å decreases from 5 to 0·05 when the absolute intensity of the 5577 Å emission increases from $5 \cdot 10^3$ to 105 rayleighs. There are indications that the atomic nitrogen emissions vary to a different degree from the atomic oxygen emissions with the increase of auroral intensity. The characteristic features in spectra of the second type represent an average between those of the first and third types.

Irrespective of the origin, the discovery of the nonlinear dependence between the total auroral intensity and the relative intensities of some atomic emissions makes it possible to give a tentative evaluation of the mean surface brightness of the most effective auroral formations under overcast sky. This is of great value for future observations of auroral displays. An additional indication of the intensity can also be obtained from the sodium and hydroxyl emission which change little during auroral phenomena.

At present there is no generally accepted explanation of such nonlinear dependence. However one cannot exclude the possibility of explaining the constancy of average relative intensities of the oxygen green line and bands of the first and second negative systems of molecular nitrogen by the mixing of the atmosphere and, consequently, by the constancy of its composition in the region of the most intense auroral emission at heights near 100 km. In this case the number of excited atoms and molecules increases as the corpuscular stream penetrates downwards due to an increase of the atmospheric density, but the relative excitation of various states undergoes no changes with height. The diminution of the 6300 Å emission at low altitudes is easily accounted for by the increase of de-activating collisions there, since the average lifetime of the initial state involved is about 100 sec. The decrease of N II emission near 5004 Å and of OI emission near 8446 Å with the decrease of height is apparently indicative of the decrease of the degree of molecular nitrogen and oxygen dissociation at lower altitudes. It hence follows that the main part of the 5577 Å emission at low altitudes cannot be attributed directly to collisions of corpuscles with oxygen atoms. If such a mechanism were

operative the intensity of this emission would grow linearly with that of the 8446 Å oxygen emission, which intensity is well correlated with that of the N II emission. This last almost certainly originates at high altitudes where the dissociation is complete. Therefore, to make the explanation agree with the picture observed, it should be supposed that 5577 Å emission at low altitudes is mainly excited in collisions of corpuscles with oxygen molecules, in which the molecules dissociate to an unexcited and an excited atom, or to two excited oxygen atoms. At the same time the emissions 5004 and 8446Å are excited by corpuscles in collisions with nitrogen and oxygen atoms respectively while the 5577 Å emission is excited this way less effectively. It should also be pointed out that the correlation between the oxygen green emission at 5577 Å and its infrared emission at 8466 Å in high altitude spectra of the first type is indicative of good mixing of the atmosphere and, consequently, the constancy of the relative composition in the region of full dissociation of atmospheric molecules at heights more than 200 km.

In connection with the question of the persistence of the auroral glow it should be pointed out that the observed afterglow exists only in the 5200 Å atomic nitrogen green emission of low intensity, in the 6300 Å emission and to a lesser extent in the 5577 Å emission. However, their afterglow is fully explained by the average lifetimes of their initial states. In no permitted intense auroral emission were traces of afterglow discovered. As is known, the mean lifetime of the recombination glow is determined by the inverse value of the product of the electron concentration and recombination coefficient. Since in aurorae the electron concentration is usually close to $10^6 cm^{-3}$ and the recombination coefficient can hardly be more than $10^{-6} cm^3 sec^{-1}$ the duration of the recombination glow is not less than 100 sec. A very important conclusion can be drawn from this, namely that all basic bright auroral emissions in the spectral region accessible to modern technique appear only as a result of impact excitement and have nothing in common with ionospheric recombination glow.

F. K. Shuiskaya[21,22] has discovered that as the latitude of the place of observation is increased the relative population of the upper vibrational levels of the $C^3\Pi$ state of a nitrogen molecule (which give the second positive system) generally increases. In addition to this G. I. Galperin and M. A. Lavrova have shown that in the case of atomic emissions the relative population of radiating atoms on upper levels also generally increases with the increase of the latitude of the place of observation.

Since in high latitudes low bright auroral formations and in low latitudes, higher auroral formations or tips of high-latitude formations are most frequently observed, it may be concluded that less energetic corpuscles predominate in higher regions, or over low latitudes in comparison with high latitudes; in other words, that the energy of corpuscles increases with the decrease of height or the increase of the latitude of the place of observation. The possibility has been mentioned of explaining the appearance of auroral-formation superimpositions extended upwards by anomalies in the angle of the pitch of corpuscles to the geomagnetic line of force without change of their rigidity. However, the observed picture does not confirm such a possibility since the upper parts of aurorae have different excitation.

By means of continuous automatic electrophotometry N. V. Jorjio[2,23] has established that time variations of the 6300 Å emission on the one hand and other emissions on the other sometimes do not coincide absolutely even for time intervals considerably exceeding the mean lifetime of the initial state for 6300 Å emission, i.e., about 100 sec. However the intensities of all other bright emissions change simultaneously even for very short time intervals: for 5577 Å they are changed for the time interval down to 1 sec and for other emissions for still shorter time intervals. During intense auroral coruscation which is most characteristic of the final stage of auroral displays separate, usually not rhythmic, bursts and fadings of emission are definitely observed with duration sometimes shorter than tenths of a second distinguished by longer time intervals. The form of the intensity variation is, in general, of com-

plex nonharmonic character. Only during very frequent coruscations does some kind of regularity appear. The conclusions may be drawn that the rapid intensity variations observed can scarcely reflect harmonic fluctuation of the geomagnetic field of the outer atmosphere. As far back as in 1958, I pointed out the possibility of hard electron generation as a result of such fluctuations of the geomagnetic field[24,25]. N. V. Jorjio's data, however, apparently show the reverse, namely, that short-period variations of this field arise from a non-uniform influx of corpuscular streams. The short-period variations of the geomagnetic field and the auroral intensity coruscations appear in the final stage of auroral development and do not carry in themselves the predominant energy of an aurora or a geomagnetic disturbance on the whole. Besides, it is very improbable that the energy of short-period geomagnetic variations can be transformed with high efficiency into kinetic energy of very fast charged particles.

Since during oscillation simultaneous changes of the intensity contour remain for emissions with initial levels with markedly different excitation energies (for example, about 4 eV for the green oxygen emission and 20 eV for bands of an ionized nitrogen molecule), they cannot originate in variations of corpuscular energy of the order of the excitation energy of the initial levels. In such a situation emission bursts from higher levels would be of shorter duration than emission bursts from lower levels which in reality is not the case.

Therefore electrons of magnetohydrodynamic waves with energies about hundredths and tenths of an eV cannot be a modulating agent, and the described auroral coruscations cannot be ascribed to them. In principle two variations of intensity modulation are possible. Firstly, the appearance is probably of considerable variable electromotive forces which retard or increase the velocity of emission-causing electrons with energies of several keV. In this case the glow modulation would be created due to change of the depth of penetration of such electrons because of some change of the effective cross-section of the interaction with atmospheric atoms and molecules. The change of the depth of penetration of corpuscles during auroral coruscations apparently is rarely observed. Since the 6300 Å emission is strongly quenched in denser layers of the atmosphere it would be possible to try to explain in such a way some counter-phase character in changes of this and other emissions. Secondly, the cause of auroral intensity modulation could be the appearance of temporal small-scale irregularities in the geomagnetic field due to magnetohydrodynamic waves or magnetic fields which are frozen inside the corpuscular streams penetrating into the Earth's atmosphere. Acting like a relay, such anomalies of short duration will increase the precipitation of the great number of geo-active corpuscles out of the Earth's outer atmosphere. In this case not only the precipitation of continuously incident corpuscular streams of moderate energy and the decrease of their repulsion in the coruscations are absent, for instance, when movable ray forms of aurorae appear; electric currents of some small rays create movable magnetic irregularities which can also lead to leakage of charged particles from more distant regions of the geomagnetic field. Sometimes the projection of the rayed arc to the Earth's surface becomes sinusoidal. Waves of such a drapery-like band move sometimes eastwards and sometimes westwards for minutes and perhaps even for many minutes. When draperies are near the zenith, which are called by coronal forms, the movable discrete electric currents of their rays will most effectively induce stable short-period electromotive force and currents below on the Earth's surface.

B. P. Potapov *et al*[26] found that absorption of 31 Mc/s cosmic radio-emission, which indicates increased ionization in the *D*-layer, begins during the appearance of intense movable ray forms in the zenith at low altitudes. The maximum absorption lags behind the maximum auroral brightness in the zenith by 1 to 10 min. Such a delay clearly shows that the corpuscles which cause an aurora, and the hard charged particles from the geomagnetic trap which cause the ionization in the *D*-layer, are not identical. Confirmation of this is provided by the absence

of such an absorption in the *D*-layer in the case of most other auroral forms excluding low movable rays.

Intensity variations of emissions which are not connected with the hydrogen emission, (i.e. emissions excited by electrons) have so far been considered. It would be of interest to compare the rapid variations of these emissions with the rapid variations of the hydrogen emission. O. L. Vaisberg has made such a comparison and found that the two do not coincide. This is of relevance in connection with the mechanism of auroral intensity modulation.

The independent behaviour of the 6300 Å emission accompanied in some cases by an increase of its intensity by several orders of magnitude is consistent with the view that it arises from the excitation of oxygen atoms by thermal electrons. T. M. Mulyarchik[27–29] using an interferometer has proved that the temperature corresponding to the width of the line at 6300 Å from the region of red forms really does sharply increase during intense aurorae reaching in some cases 3500°K whereas in normal aurorae it is approximately 1000°K. The increase of the temperature in the auroral region is found also from the width of emissions which originate at lower altitudes, but in this case it does not reach the values which are characteristic of the region from which 6300 Å originates. However, calculations have shown that even the increase of the temperature of the outer atmosphere indicated is not sufficient to explain the observed increase of the 6300 Å intensity. G. I. Galperin believes that this indicates that the temperature of the electrons is higher than the temperature of the neutral oxygen atoms, i.e., that there is not thermodynamic equilibrium.

The proof of great heating during auroral displays is of paramount importance for the physics of the upper atmosphere. With the temperatures mentioned loss of neutral atoms and molecules from the atmosphere is inevitable. Besides, with such temperatures ionized particles will penetrate along the whole length of the geo-magnetic lines of force and form dense filaments of an ionized gas. This creates an unexpected situation in the exosphere and new conditions for magnetohydrodynamic wave propagation and impingement to the Earth of various corpuscles and outer magnetic fields. The increase of the scale height during auroral displays explains the existence of greatly extended rays having quite a distribution of luminosity with altitude.

Up till now only aurorae which are not in the dark part of the atmosphere have been considered. However, when they are lit by solar radiation, additional very intense fluorescent emissions appear. Like many other investigators, F. K. Shuiskaya[21,22] determined the vibrational temperature of fluorescent bands of the first negative system of molecular nitrogen. In harmony with the interferometric measurements she recorded very high temperature in the upper atmosphere during auroral displays.

When aurorae are illuminated by the Sun, especially over low-latitude regions, a movement upwards of fluorescent ionized nitrogen molecules is very graphically observed, at first in the form of separate narrow filaments and then of more vast formations. This is a direct proof of the expanding of the upper atmosphere during auroral phenomena.

However, the most interesting and sensational discovery was the presence of the 10,830 Å helium infrared emission with intensity of more than 10^4 rayleighs in sunlit red aurorae. The occurrence of this emission without an explanation of its nature was reported in 1958 in a brief note by A. V. Mironov, V. S. Prokudina and N. N. Shefov[10]. At first, we viewed with great caution identification of the feature at 10,830 Å with the infrared helium lines. It was difficult to see how these lines could be excited. Numerous systematic observations by N. I. Fedorova[30,31] at Loparskaya during subsequent years exclude any doubts in the reality of the 10,830 Å emission in sunlit red aurorae. N. N. Shefov[32,33] has recently explained it as being a fluorescent scattering of solar radiation by helium atoms in the metastable 2^3S state. A detailed analysis has shown that the metastable atoms are produced by collisions of ~25 eV electrons with normal helium atoms. Hence in the region of red aurorae in addition to the

many electrons with energies of about 2 eV which must exist in order to excite the oxygen red emission, there must be many electrons with energies of about 25 eV since the helium concentration cannot be high. It is quite possible that the electrons with energies of about 25 eV form an independent group of geo-active corpuscles which causes red aurorae. However, such electrons can also appear in the upper atmosphere as a result of its ionization by streams of electrons with energies of about 10 keV discovered by us with the aid of Sputnik III[21, 25, 34]. Thus, observation of the 10,830 Å emission opens up new possibilities of regular investigation of the upper atmospheric helium and also of geoactive electrons. N. N. Shefov has developed a comprehensive theory of the helium emission. He predicts that it should be observed regularly in the twilight glow though with rather lower intensity. In this case the singlet states are first populated by the solar emissions at 537·1 and 584·4 Å. Some of the excited helium atoms cascade to the 2^1S state from which they are transferred into 2^3S state in encounters with thermal electrons. N. N. Shefov has proved his prediction by direct observations[35]. The intensity of the helium twilight emission is about 10^3 rayleighs; thus it is one of the strongest that occur. By regular recording the fringes of this emission in twilight with the aid of photocontact tubes, narrow-band filters and interferometers, P. V. Shcheglov[36] has provided an additional confirmation of Shefov's prediction. Since this twilight feature reflects the intensity of the solar helium lines at 537·1 and 584·4 Å, it is difficult to overestimate the scientific and practical importance of the remarkable discovery for the studies of the upper atmosphere and of solar activity.

Bright fine auroral rays sometimes acquire a red-orange tint on one of the sides even near the base. The reality of such coloration is not in doubt. Nevertheless even with modern techniques we failed to find any difference in the spectra of such formations and the spectra of the usual rays with a whitish-green tint. In our report at the Congress of the Union of Geodesy and Geophysics in Toronto in 1957 this phenomenon was described as a "wind-vane effect"[37]. The phenomenon appears when there is a relative motion of a corpuscular beam and the atmosphere. It does not matter whether the beam or the atmosphere is at rest. When the air is bombarded by corpuscles, the high transition probability emissions of the first positive system of molecular nitrogen which have a red-orange tint (the mean lifetime of the initial state is about 10^{-6} sec) begin to glow in the visible region. However, the glow of the low transition probability 5577 Å oxygen green emission (the mean lifetime of the initial state is about 1 sec) lags behind by about 1 sec. Therefore the part of the atmosphere which is crossing the beam of corpuscles has a red-orange glow and parts which have already crossed it have a greenish glow. Due to a small interval between the sides of a movable ray, which is not resolved by the height of a slit with our modern spectral instruments, different emissions are superimposed. For instance, with usual velocities of the atmosphere relative to the beam of corpuscles of hundreds m/sec the distance between differently colored parts of a fine ray is approximately 100 m. We have repeatedly observed a chaotic play of ray colours which testified to a very complicated and intricate circulation of the upper atmosphere, possibly even of eddy motions. The phenomenon of a "wind-vane effect" is clearly observed only when a very definite focusing of separate small rays takes place. It is difficult to observe in the case of formations without sharply outlined boundaries. When the angular resolution and sensitivity of the registering spectral instruments are improved, a "wind-vane effect" will open rich possibilities for investigating the circulation of the upper atmosphere since the velocity of the motion of the corpuscular beam relative to the observer can be determined and can then be taken into account when the actual motion of the upper atmosphere is determined. Of special interest is the observation of a "wind-vane effect" not only in the green, but also in the emission at 6300 Å or in the nitrogen green emission at 5200 Å, the mean lifetime of whose initial states is much greater. In this case, the separa-

tion of different emissions will be observed at considerably greater distances.

The study of auroral photographs which are at our disposal has revealed very strange formations (see Fig. 1). It seems that they originated during the piercing of clouds consisting of a medium which is more effectively excited by beams of corpuscles than is the surrounding atmosphere. Since the weather during the aurora described was cloudless, the reality of the structure photographed is beyond any doubt. In the collection of our photographs we have also found many in which a sharp intensity jump occurs in the low part of all rays along the common boundary (see Fig. 2). The aurorae concerned were completely in the shadow of the Earth.

The auroral forms described can be explained by a chaotic or organized circulation of some layers of the upper atmosphere due to which the strongly dissociated air with atomic nitrogen, in which intense molecular emissions cannot originate, moves downward. Intense molecular emissions appear in clouds or in the low air where the abundance of molecular nitrogen is great. Since the scale of the intensity changes is much less than the scale height, the corresponding great gradients in the allotropic composition can be preserved only during an intense circulation and are impossible in a quiet atmosphere. Further observation of such aurorae seems to extend the possibility of investigating processes of circulation and mixing of the upper atmosphere in the auroral zone.

Having analyzed phenomena associated with the upper atmospheric circulation one should note that they do not play a major part in most aurorae. Apparently, the circulation in the upper atmosphere in the region of aurorae created by corpuscles is not great. Otherwise it would be easily evident from the intense "wind-vane effect." In principle acceleration of charged particles is possibly due to dynamo-effect outside the regions with corpuscular streams. (in low latitudes for instance) if in this process only those particles appear whose energies do not exceed the excitation energy of the initial states for the emission of the first and second positive and negative systems of molecular nitrogen. Thus, it would be possible to try to explain the intensity enhancement of some characteristic emissions in auroral spectra of the zeroth and first types without traces of emissions of nitrogen molecular bands. The presence in the upper atmosphere of electrons with energies exceeding 7 eV could not have remained unnoticed because of these characteristic emissions (which are usually absent in the night airglow). The nature of red arcs in equatorial regions is still obscure. Since magnetic anomalies in the region of the display can lead to the release of high-energy charged particles from the geomagnetic trap, their discovery over these arcs by means of artificial Earth satellites is not a conclusive argument against the hypothesis of the creation of red equatorial arcs by dynamo effect.

Now let us recall another phenomenon noticed at Loparskaya during the rapid projection of ascafilms. It had already been described briefly in 1957 before the beginning of the I.G.Y.[38] and in our report at the Congress of the International Union of Geodsey and Geophysics in Toronto in 1957[37]: "In previous years visual observations at Loparskaya made it possible to suppose the existence of large-scale curling motions of auroral luminous formations. At the end of 1956 this superimposition was confirmed by means of films taken by all-sky cameras. It can even be said that by the end of 1956 a different picture of the auroral development has not yet been observed. The development of luminous formations starting usually from the quiet arc takes place along the ellipsoidal spiral, the major axis of which is extended by hundreds and thousands of kilometres. In ellipses extended along the magnetic parallel the minor axis then increases. In the last stages rayed forms are developed which are transformed into diffuse pulsating aurorae."

Sometimes the structure and brightness of the aurora undergo such rapid and irregular changes that concentric motions can be observed only during very rapid projection of ascafilms. Motions of light formations are observed most graphically in rayed arcs with fine small "focused" rays. Each of these small rays lights up, and is quenched. The sharper

the focusing of separate small rays the more definite is the picture. Motions during magnetic storms are of a very complicated and intricate character. A more definite picture is traced in aurorae of short duration on magnetically quiet days. In rayed loops, in special and more complicated formations, motions take place either clockwise or in the opposite direction. Sometimes several such formations are observed simultaneously with the concentric motions in different directions. It is quite probable that draperies are one of the forms of motions along the broken trajectory. Sometimes motions occur along a very flattened spiral. When the loop of such a spiral is beyond the horizon, two parts of one looped rayed arc can be taken for independent formations. In southern regions, for instance, at Zvenigorod, motions along a spiral were not observed. However, here too this motion was in the western as well as in the eastern directions. It is very interesting to note that the centres of some auroral formations with concentric motions remain in a constant position relative to the Earth for a long time. Here of course only sharply outlined auroral forms have been meant. Diffuse formations against whose background these forms appear have a more continuous structure along the longitude.

Since the main observed auroral emissions are produced by the streams of electrons of moderate energy from the outer parts of the atmosphere, it is possible that the drift of light formations reflects the drift of these streams in the space and time-varying geomagnetic field. The drift velocities, which are equally probable either in eastern or in western directions, reach some kilometres/sec if the aurora is at a height of 100 km. These velocities are greater than the maximum velocities of any wind motions possible in the upper atmosphere. If such drifting particles pass through the geomagnetic equatorial plane, the drift velocity will naturally be greater in the plane.

If this is really the drift of not very energetic electrons causing aurorae then it must be stated that it does not, as has been assumed lately, follow the magnetic parallels in an easterly direction in the same way as the drift of energetic electrons in the so-called "radiation belts." In reality there is no definite direction of the drift. The observed picture is most easily accounted for by the drift of electrons which cause aurorae in arbitrary irregularities of the outer part of the space and time-varying geomagnetic field among which there are cavities with increased and decreased intensities. Only in this case is the drift in different directions possible. Since the energy of electrons which cause aurorae is about 10 keV and lower, even in an undisturbed symmetrical geomagnetic field it will take almost 10 or more days for electrons located on magnetic lines of force of geomagnetic latitude of 67° to go once around the Earth. It is quite evident that the mean shift to the east for several hours will amount to several degrees or lower. In view of this the low mobility of the centres of the concentric drift becomes quite understandable. Electrons which cause some auroral formation fall out much more quickly than they can carry out a considerable drift motion in the geomagnetic field. It is quite evident that such electrons cannot create a concentric zone about the Earth. During their existence they will form only filaments along the geomagnetic line of force. Therefore there is nothing surprising in the fact that aurorae have strongly changeable longitudinal intensity and structure.

The longitudinal uniformity of the hydrogen emission deserves special consideration. A suspicion arises that this uniformity cannot be explained only by imperfection of observations and by diffuse uncertainty of the hydrogen emission fields due to change exchange. For final elucidation of this phenomenon further more precise observations are required. However, it is quite probable that the uniformity is connected with higher drift velocity of primary protons, if they originally have higher energy than radiating hydrogen atoms observed. A higher Larmour radius of such primary protons could have also ensured lower dependence of their drift on irregularities of the magnetic field.

From the above considerations it is evident that "radiation belts" can appear only due to swiftly drifting high-energy charged particles. However, these particles do not play an essen-

Fig. 1

Fig. 2

(*To face page* 136)

Fig. 3

Fig. 4

Fig. 5

tial part in an auroral display since their total power is insignificant.

The size of the magnetic field irregularities determine the drift of the electrons which cause aurorae and can be evaluated from the dimensions of the auroral loops shown in Figs. 3 and 4. In the equatorial plane, where it intersects with the geomagnetic lines of force from the auroral zone the size of irregularities will be much higher than that of the loops observed—it will be thousands of kilometres. However, a very high velocity of the drift, which exceeds by many times the velocity of the drift of the same electrons in an undisturbed geomagnetic field, indicates that the gradients of the irregularities of the magnetic field considerably exceed those of a normal geomagnetic field. As is known[39], the drift velocity v in a non-homogeneous magnetic field in a cm/sec can be written as

$$v \sim E \times \frac{c}{e} \times \frac{\nabla}{H^2}$$

where E is the energy of an electron in ergs, c is the velocity of light in cm/sec e is the charge of an electron in CGSE, H is the intensity of a magnetic field in Oe CJSM and ∇ is the gradient of the same value in the same system of unit. For $E \sim 10$ keV when in the equatorial plane $v \sim 10^9$ cm/sec and $H \sim 10^{-3}$Oe, $\nabla \sim 10^{-12}$Oe cm^{-1}. With an irregularity of 1,000 km the total change of the magnetic field intensity will be 10^{-4}Oe.

Magnetic fields of such intensity which were expected earlier[40], have already been recorded in solar corpuscular streams[41]. Therefore the appearance of the irregularities mentioned above can be associated, as we have already suggested, with magnetic fields frozen inside the streams of primary corpuscles[24, 25]. These fields can also carry from the Sun more energetic electrons and protons than those of the main mass of the particles in these streams. Agglomerations of corpuscles cemented by the magnetic field apparently can explain their deep penetration into the equatorial plane near which aurorae are now observed and where at high altitudes direct methods led to the discovery, not only of hard electrons, but also of electrons with the energy needed to produce aurorae. Such formations can have in the geomagnetic field a very great penetrating capability since the ratio of their total charge or magnetic moment to their total mass can be much lower than that of individual charged corpuscles. When this agglomeration of corpuscles strikes a geomagnetic field and again in the outer atmosphere around it, a redistribution of kinetic and potential energies of the agglomeration will inevitably take place between all interacting elements. The energy of the primary corpuscles will be charged and many ions of the Earth's atmosphere will acquire high energies, i.e., they themselves will become geo-active corpuscles.

Irregularities of a geomagnetic field which originate in the above described processes can hardly be symmetrical in respect of an equatorial plane. Therefore a detailed correspondence of aurorae borealis and aurorae australis is excluded in principle. Under these conditions electrons which cause aurorae need not necessarily (before their full precipitation) oscillate along the line of force of the geomagnetic field from one polar region to the other. Such is apparently the situation in reality. The region with an unhomogeneous magnetic field does not necessarily belong to the Earth's outer atmosphere.

The energy of electrons which cause normal low aurorae is close to 10 keV. Strong currents of electrons with such energies have been discovered by us at very high altitudes over the zone of normal aurorae[24, 25, 34]. It is remarkable and scarcely accidental that the energy of the protons of the primary corpuscular streams is also close to 10 keV. On the one hand, this may mean that the protons and electrons of corpuscular streams penetrating into the atmosphere have nearly equal temperatures; on the other hand, it is possible that as a result of some interaction in the outer geomagnetic field the protons lose the predominant part of their energy, imparting it almost completely to the electrons. It is very interesting that, from the north and the south, fields with emission of hydrogen atoms with energies of several hundreds eV are limited by emission fields excited only by electrons with energies of about 10 keV.

It is not impossible that this process is connected with the formation in the geomagnetic field of potential gradients which ensure the transference of energy from the primary protons to the electrons of the outer atmosphere. Lastly, during induction processes linked up with the change of magnetic fields in interplanetary space the appearance of protons and electrons with equal energies can also be expected in the outer atmosphere. However, in this case the induced electro-motive force will not have some definite limited value of about 10 keV and charged particles will be accelerated up to very varying energies greater or less than 10 keV depending on the rate and amplitude of the outer magnetic field charge.

In connection with these considerations the following circumstance is of importance. Fine structural ray formations are characteristic of movable forms of aurorae (rayed arcs, draperies). A very fine structure of small rays sometimes hundreds of metres in cross-section indicates that the electrons which cause them are nearly monochromatic. Small rays would be de-focused if the velocity of the exciting electrons had a dispersion, and various drift velocities would be caused. With observed drift velocities of light formations in the auroral zone from zero to a km/sec and lifetimes of individual small rays of some tenths of a second, their width would amount to several kilometres. Such a short-period existence of individual small rays may be explained as follows. Electron agglomerations, which oscillate along the geomagnetic lines of force, have a length which is much shorter than the outer lines. The mechanism of monochromatization of the electron velocities is not clear. To outline the principle suggestion in this direction let us note that rayed forms usually appear in the main phase of a geomagnetic disturbance simultaneously with long harmonic short-period variations of a geomagnetic field which, as was already pointed out, are like a relay, capable of governing the precipitation and the repulsion of electrons. It is quite possible that during oscillations along lines of force electrons are selected with a quite definite velocity which can then in addition be subjected to self-focusing in the electro-conducting outer atmosphere according to the mechanism suggested by Bennet and Hulburt[42]. Most of the electrons remain. They have a considerable dispersion of velocities and ensure the excitement of the more intense diffuse auroral forms. Figure 5 shows the beginning of the change of direction of the arc against the background of diffuse light accompanied by the surrounding of diffuse formations in the north by a rayed arc of a very fine structure. Quiet homogeneous forms during minimum solar activity may be explained by the fact that solar corpuscular streams at this time are more uniform, do carry strong enough magnetic fields frozen inside them and, therefore, are not accompained by intense short-period geomagnetic field variations or some other processes which cause monochromatization of velocities of electrons causing aurorae. Here there is much room for further investigations.

A great gap in our knowledge of aurorae is caused by the absence of day observations. Having been engaged in optical investigations of aurora we hoped that radar would give at least a statistical picture of such phenomena. Great was our surprise when it turned out that radar echoes from aurorae during daytime are practically absent in the form in which they are observed during night time[43]. It would be premature to conclude that this proves a real absence of aurorae during daytime: further investigations are needed. Perhaps the disappearance of radar echoes is explained by the change of location of an auroral zone in which the geometry of irregularities of ionization favourable for the echoes is disturbed. This is quite possibly due to a shift by day of the auroral zone closer to the pole, i.e. a shift contrary to that at night.

All auroral problems cannot be solved by ground observations. A daytime picture of aurorae, the detailed structures and location of the streams of protons and electrons which cause aurorae, their origin, the transformation of the primary corpuscular streams, their initial composition, their nonhomogeneity and many other things can be established only by studies in the Earth's outer atmosphere and in

interplanetary space. Investigation of the uniformity of the drift-current ring about the Earth is of special interest. Researches in this direction are now only in their early stages of development.

THE NIGHT AIRGLOW

In beginning to discuss new results in connection with the night airglow it should be emphasized once more that there is some uncertainty as to what refers only to auroral phenomena and what only to the night airglow. This uncertainty is especially great when we deal with oxygen red emission at 6300 Å and hydrogen emission H_α.

As already reported our Institute has during recent years regularly recorded hydrogen emission with a narrow contour the width not exceeding 2 Å and being usually somewhat less. This prevents us from supposing that the radiating hydrogen atoms have velocities greater than 10^7 cm/sec and energies of greater than some tens of eV.

N. N. Shefov in our Institute and V. I. Yarin in Yakutsk have acquired rich collections of nightglow spectra from 3000 to 12,000 Å[44-52]. They have observed lines of OH, O_2 and OI. The total number of permitted lines of O_2 known is now 500 which is very many times greater than the number known earlier. A night airglow continuum is also recorded. It has become evident that during electrophotometry of the night airglow this continuum should be very carefully taken into account in order to prevent distortion of the true values of discrete emission intensities investigated. In the blue spectral region continuum intensity bursts have been discovered. Simultaneously strong Herzberg bands of molecular oxygen appear. At present, however, any correlation of this emission with other emissions of the nightglow has not been found.

Processing observational data on the hydroxyl emission at Zvenigorod has shown that the rotational temperature varies from 200 to 400°K. Sometimes temperature changes of some tens of degrees occur from night to night. They cannot be attributed to real temperature changes of the atmosphere below 100 km since this would require an impossibly large amount of energy to be absorbed. On the other hand, we have no doubts that the hydroxyl rotational temperature reflects the temperature of the surrounding medium. Therefore the great temperature changes observed from night to night can be accounted for only by the change of the height of the radiating layers. The multilayer character of the hydroxyl emission was proved by temperature differences in bands from different initial vibrational levels. It is noteworthy that there were variations in the relative populations of different initial levels. This can be explained only by the existence of at least two excitation mechanisms.

V. I. Yarin in Yakutsk obtained a similar result in general. However while in Zvenigorod no interdependence between the intensity of the hydroxyl bands and their rotational temperature was found, in Yakutsk the intensity tends to be an increasing function of the rotational temperature when this is above 250°K. The observed increase of the hydroxyl emission may be due to the ozone-hydrogen reaction with activation energy of about 3 kcal per mole. In Yakutsk seasonal intensity and temperature variations of hydroxyl bands are observed. Their maximum value is greater in the cold winter season. In Zvenigorod such dependence was not observed. It therefore appears that the hydroxyl emission peculiarities depend on the place of observation.

N. N. Shefov and V. I. Yarin[53] compared all the data at our disposal on the temperature of the hydroxyl rotational bands and confirmed that this temperature tends to increase with the latitude of the place of the emission inceptor. In Zvenigorod the registered intensity of the hydroxyl bands does not decrease appreciably even in extremely overcast weather, a fact which presumably is indicative of the increase of their actual intensity at this time. Average intensities of all the hydroxyl bands from 5000 to 12,500 Å have been obtained. It follows from them that the rate of formation of excited hydroxyl radicles is about 10^{12} per cm vertical column per sec. The hydroxyl band from the tenth vibrational level is also recorded.

N. N. Shefov, V. I. Yarin and V. S. Proku-

dina[18,19,47,59] have established that the night sky emissions can be sub-divided into two groups. Atomic oxygen emission at 5577 Å, belongs to the first group. Hydroxyl emission, sodium emission at 5894 Å, H_α hydrogen emission and atomic oxygen emission at 6300 Å belong to the second mutually correlating group. The groups are not co-variant. This correlation breaks down when galactic fields with hydrogen emission get into the spectrograph field and when the intensity of 6300 Å becomes high. In all cases there is clear correlation between the intensities of 6300 Å and of the green nitrogen emission at 5200 Å. A. V. Mironov[54] has discovered an association between the enhancement of these emissions and the magnetic K-index.

A correlation between the hydroxyl and sodium emissions, on the one hand and the hydrogen and red oxygen emissions, on the other is of great interest. These pairs of emissions clearly originate at different altitudes. The first pair originates below 100 km, the second pair originates much higher. Their correlation may perhaps be explained in the following way.

Red oxygen and hydrogen emission are due to the same exciting agent which acts high in the atmosphere. The correlation of these emissions with emissions from low in the atmosphere arises from this agent directly through intermediate products, regulating the concentration of atomic hydrogen and hence the hydroxyl and sodium emissions. Electrons with energies of about 10 keV or hydrogen L_α radiation might be the primary agent involved. X-rays produced during bombardment of the atmosphere by energetic electrons and hydrogen L_α radiation could penetrate downwards and cause dissociation of hydrogen molecules[25].

CONCLUDING REMARKS

In the course of our analysis we have repeatedly compared the intensity of an emission with geomagnetic characteristics. On the average a statistical agreement is observed. It seems quite evident, especially when aurorae are considered. However, in many cases there are undoubted disagreements, and there is nothing surprising in this also. By means of emissions, elementary processes are observed in limited volumes of the upper atmosphere while geomagnetic phenomena and macroscopic ionospheric characteristics reflect the integral effect of vast space even beyond the visibility range from the place of observation of emissions. As a result complete and constant coincidence of circumstances would be doubtful. An advantage of geomagnetic and ionospheric observations is their independence on time of day or meteorological conditions. Detailed information on the primary cause of phenomena is however difficult to obtain from them. Observations of upper atmospheric emissions give direct information on primary elementary processes which lead to the formation of the ionosphere and to geomagnetic disturbances. Unfortunately, such observations are possible now only at night-time under good meteorological conditions. It is much more difficult to find a method of ground observations of upper atmospheric emissions during the day. However, direct sounding of the upper atmosphere considerably extends the scope of possible investigations of the elementary processes occurring. It seems likely that in the immediate future many problems which are not now clear will be solved.

REFERENCES

1. *Sbornik*,* N **1**, 5, 1959.
2. N. V. Jorjio, *Sbornik* N **1**, 30 1959.
3. D. Barbier, *Ann. Géophys.* **16**, 544 (1960).
4. B. J. O'Brien, J. A. van Allen, P. E. Roach, C. W. Gartlein, *J. Geophys. Res.* **65**, 2759 (1960).
5. V. S. Prokudina, *Sbornik* N **1**, 43, 1959.
6. G. I. Galperin, *Astr. J. (Moscow)*, **34**, 131 (1957).
7. G. I. Galperin, *Astr. J. (Moscow)*, **35**, 382 (1958).
8. G. I. Galperin, *Sbornik* N **1**, 7 (1959).
9. G. I. Galperin, *Planet. Space Sci.*, **1**, 57 (1959).
10. A. V. Mironov, V. S. Prokudina, N. N. Shefov, *Sbornik* N **1**, 20 (1959).
11. F. K. Shuiskaya, *Sbornik N* **1**, 45 (1959).
12. A. B. Meinel, *Mem. Soc. Roy. Sci. Liège*, **12**, 203 (1952).
13. O. L. Vaisberg, *Izv. Akad. Nauk SSSR, ser. geofiz.*, N **8**, 1277 (1960).

* In this bibliography *Sbornik* means a collection "Spectral, Electrophotometric and Radar Investigations of Aurorae and Night Airglow" published by the U.S.S.R. Academy of Sciences.

14. O. L. Vaisberg, *Izv. Akad. Nauk SSSR, ser. geofiz.* N **1**, 166 (1961).
15. I. S. Shklovskii, *Dokl. Akad. Nauk SSSR*, **81**, 367 (1951).
16. I. S. Shklovskii, *Ann. Géophys.*, **14**, 414 (1958).
17. C. Y. Fan, *Airglow and Aurorae* (Ed. Armstrong and Dalgarno), p. 276. Pergamon Press, London (1956).
18. V. S. Prokudina, *Sbornik* N **2-3**, 68 (1960).
19. V. S. Prokudina, *Sbornik* N **5**, 32 (1961).
20. N. V. Jorjio, *Sbornik* N **2-3**, 45 (1960).
21. F. K. Shuiskaya, *Izv. Akad. Nauk SSSR, ser. geofiz.*, *N* **3**, 510, 1960.
22. F. K. Shuiskaya *Sbornik N* **5**, 49, 1961.
23. N. V. Jorjio, *Izv. Akad. Nauk SSSR, ser. geofiz.*, N **5**, 714, 1960.
24. V. I. Krasovskii, Report to the Symposium, Fifth General Assembly of C.S.A.G.I. also *Isk. Sput. Zemli*, USSR Academy of Sciences, N **2**, 59 (1958).
25. V. I. Krasovskii, Int. Astr. Union, **10**, 716 (1960).
26. B. P. Potapov, Z. Ts. Rappoport, T. B. Borsuk, *Sbornik* N **2-3**, 42 (1960).
27. T. M. Mulyarchik, *Sbornik* N **1**, 41 (1959).
28. T. M. Mulyarchik, *Dokl. Akad. Nauk. SSSR*, **130**, 303, 1960.
29. T. M. Mulyarchik, *Izv. Akad. Nauk SSSR, ser. geofiz.*, N **3**, 449 (1960).
30. N. I. Fedorova, *Sbornik* N **5**, 42 (1961).
31. N. I. Fedorova, *Planet. Space Sci.*, **5**, 75, 1961.
32. N. N. Shefov, *Sbornik* N **5**, 47, 1961.
33. N. N. Shefov. *Planet Space Sci.*, **5**, 75, 1961.
34. V. I. Krasovskii, I. S. Shklovskii, Iu. I. Galperin, E. M. Svetlitskii, Iu. M. Kushnir, G. A. Borodovskii, *Isk. Sput. Zemli.* U.S.S.R. Academy of Sciences, **N6**, 113 (1961).
35. N. N. Shefov, *Astr. Circ. Acad. Sci.* N 222 (1961).
36. P. V. Shcheglov, *Astr. J. (Moscow)* **38**, In press (1961).
37. V. I. Krasovskii *et al. Ann. Géophys.*, **14**, 346, 1958.
38. V. I. Krasovskii, *Priroda (Nature)*, N **5**, 55 (1957).
39. L. Spitzer, *Physics of Fully Ionized Gases*. New York-London (1956).
40. V. I. Krasovskii, *Planet. Space. Sci.*, **1**, 14 (1959).
41. L. Coleman *et al.*, *Phys. Rev. Letters* **5**, 43, 1960.
42. V. H. Bennet, D. E. Hulburt, *Phys. Rev.*, **91**, 1562 (1953).
43. B. A. Bargaryatsky, *Sbornik N* **2-3**, 11, 1960.
44. N. N. Shefov, *Sbornik* N **1**, 25 (1959).
45. N. N. Shefov, *Sbornik* N **2-3**, 57 (1960).
46. N. N. Shefov, *Sbornik* N **5**, 5 (1961).
47. N. N. Shefov, *Sbornik* N **5**, 18 (1961).
48. N. N. Shefov, *Sbornik* N **5**, 39 (1961).
49. N. N. Shefov, *Sbornik* N **6**, 21 (1961).
50. V. I. Yarin, *Sbornik* N **5**, 53 (1961).
51. V. I. Krasovskii, *Sbornik* N **5**, 53 (1961).
52. V. I. Krassovskii, N. N. Shefov and V. I. Yarin, *J. Atmos. Terr. Phys.*, **21**, 46, 1961.
53. N. N. Shefov and V. I. Yarin, *Sbornik* N **5**, 25 (1961).
54. A. V. Mironov, *Sbornik* N **2-3**, 66 (1960).

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 142–150. Printed in Great Britain.

CIRCULAR SYMMETRY IN THE GEOMAGNETIC PLANE FOR AURORAL PHENOMENA

BENGT HULTQVIST
Kiruna Geophysical Observatory, Kiruna C, Sweden

(*Received* 15 *September* 1961)

Abstract—Published observational data concerning the configuration and location of the northen and southern auroral zones are reviewed and are compared with circles projected from the geomagnetic equatorial plane, outside the earth, on to the earth's surface along the geomagnetic field lines (in an approximation which includes the five first terms of the spherical harmonic development). Other auroral distance parameters are also compared with these circle projections.

Very good agreement has been found for the northern hemisphere. More accurate data are still needed for those parts of the southern auroral zone which lie over the ocean at a great distance from the Antarctic continent. In the author's opinion the new data confirm that the northern and southern auroral zones are geomagnetically conjugated curves and that they correspond to projections of circles in the geomagnetic equatorial plane on to the earth's surface along the real geomagnetic field lines.

INTRODUCTION

There are good reasons for postulating circular symmetry, in the geomagnetic equatorial plane outside the earth's surface, of distance parameters representing statistical averages for corresponding variables of those processes, in and near the equatorial plane, which are associated with the occurrence of auroral phenomena in or near the two auroral zones. Only the dipole field affords a measurable contribution in the equatorial plane at a distance of several earth radii from the earth's surface, and the dipole field is circularly symmetric. Thus the configuration in the geomagnetic equatorial plane will not be influenced by the earth's rotation, provided minor effects of the different directions of the geomagnetic and the rotation axes are disregarded. The equatorial plane process will therefore be seen in the same distance range from all meridians on the earth. It follows that—if the processes in the equatorial plane and in the auroral zones are connected by the real magnetic field lines, as seems physically plausible—the only longitudinal dependence of e.g. the averages of a great many auroral locations observed at fixed points on the earth and equally distributed with regard to geomagnetic time, will be that introduced by the deviation of the earth's magnetic field from a dipole field. This deviation is appreciable at and near to the earth's surface.

According to this view the configuration in one hemisphere will correspond to the projection of the corresponding one in the other hemisphere along the geomagnetic field lines.

The only assumption made in the foregoing is that an electron or an ion moves from the equatorial plane to one of the auroral zones along a geomagnetic field line or on the surface defined by those field lines which cross the equatorial plane along one and the same circle and which contain the location of the particle in the equatorial plane. Such an assumption seems to be well founded. No other details of the actual physical processes are involved.

The projections of circles in the equatorial plane to the earth's surface along the real field lines, were calculated by Hultqvist[1] for purposes of comparison with experimental isochasms.

In this paper the circle projections are compared with various published observational data and with material obtained during I.G.Y.

REVIEW OF OBSERVATIONAL RESULTS CONCERNING THE LOCATIONS OF THE AURORAL ZONES

The northern auroral zone delineation of Fritz[2], and Vestine's[2] modification of it are shown in Fig. 1 together with a zone determined by Feldstein[4] on the basis of photographic recordings taken at 39 Arctic stations and visual

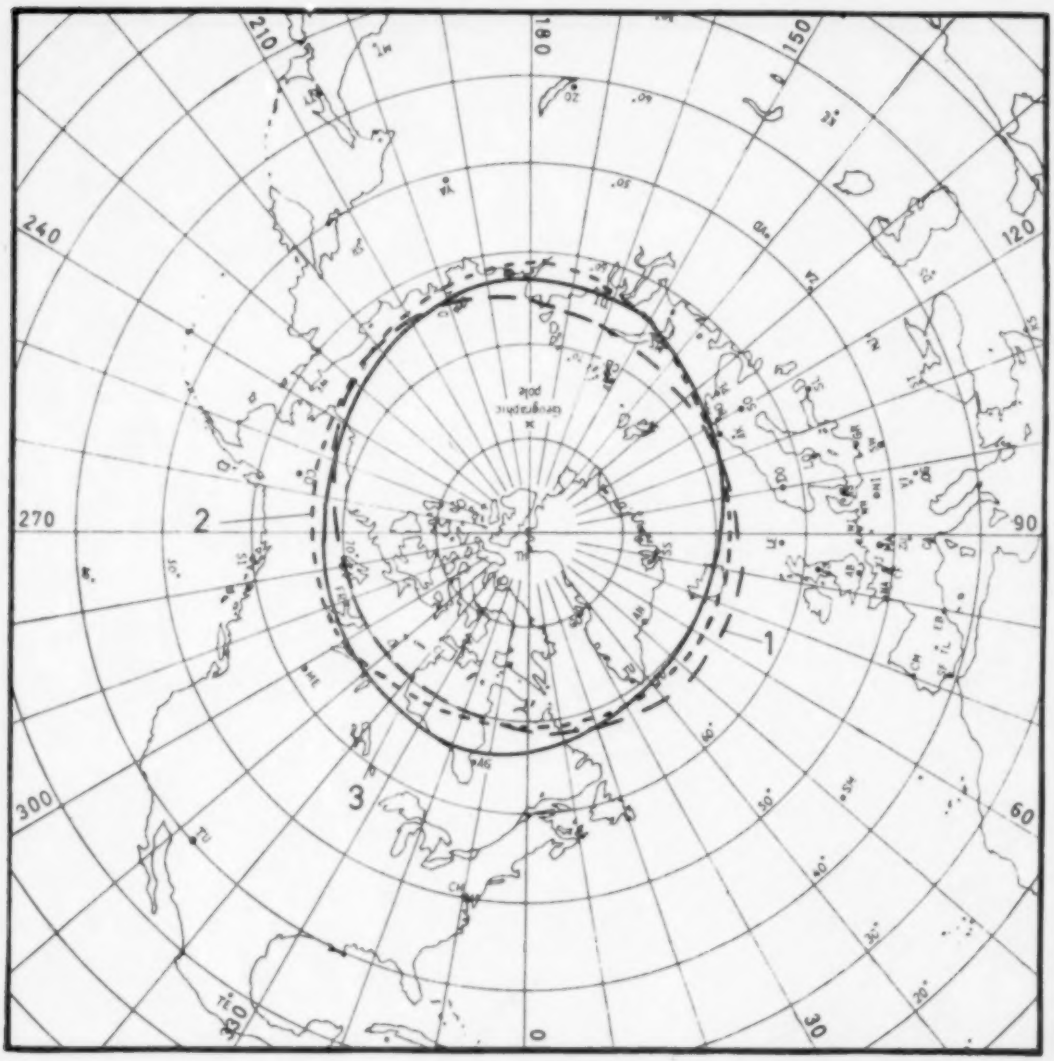


Fig. 1. Northern auroral zones determined from observations (on a map with geomagnetic co-ordinate system). Curve 1 is that of Fritz[2], curve 2 is Vestine's[3] and curve 3 Feldstein's[4].

observations made at three sites during the I.G.Y., 1957-58.

While Fritz's auroral zone was based almost entirely on unsystematic auroral observations from several hundred years B.C. up to about 1878, Vestine had at his disposal more systematic observations, particularly those from the two polar years 1882-3 and 1932-3. Both delineations, however, are based on material obtained over very long periods of time with secularly changing geomagnetic fields, varying solar activity, partly unknown meteorological conditions, etc. In contrast, Feldstein's[4] curve is based on observations made within a few years with the use of a uniform technique at all stations. The auroral indices employed in his

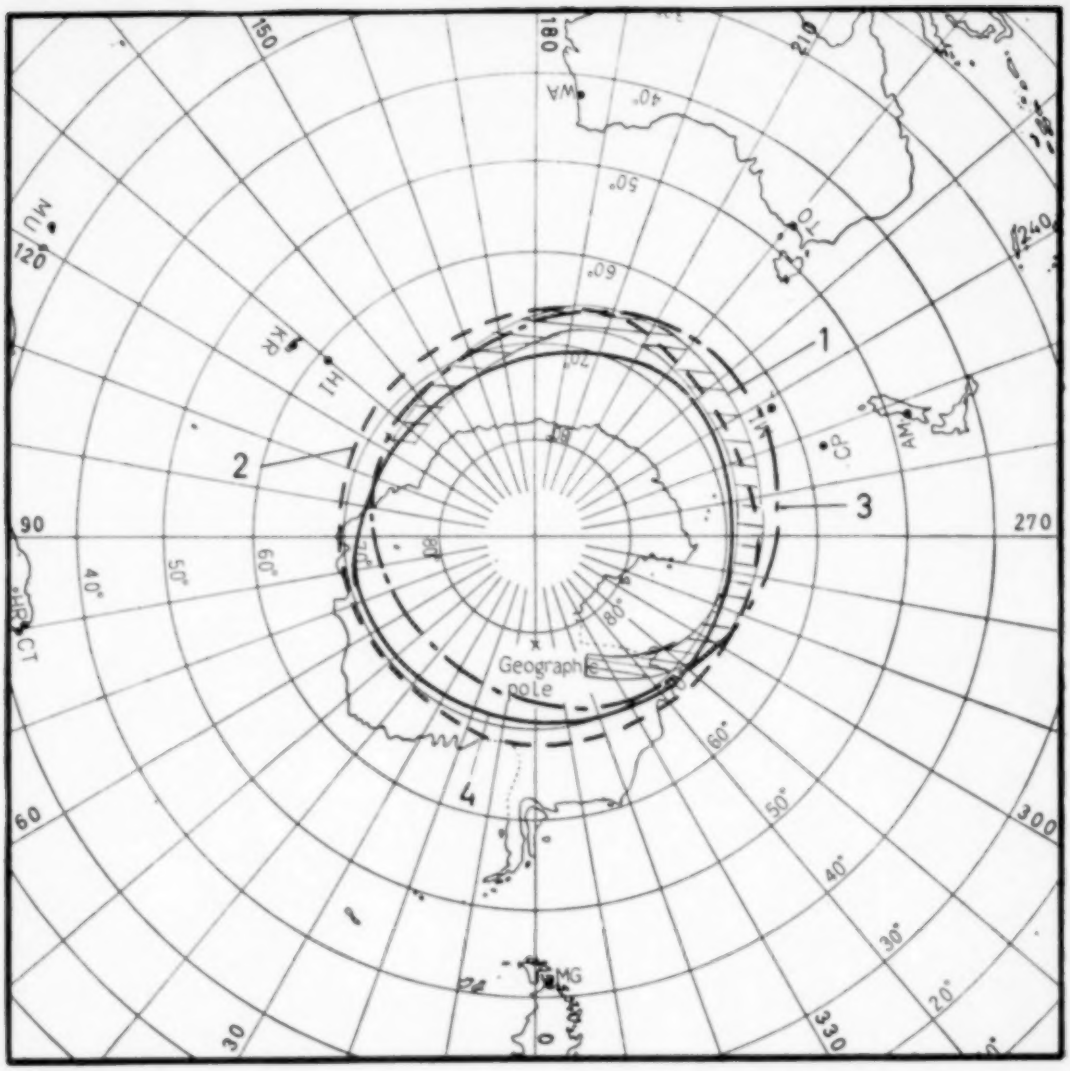


Fig. 2. Southern auroral zones determined from observations (on a map with geomagnetic co-ordinate system). Zone 1 is according to White and Geddes[7], zone 2 to Vestine and Snyder[8], zone 3 to Bond and Jacka[15], and zone 4 to Feldstein[4].

determination are more accurate than those previously used, as only the aurora in the zenith has been utilized.

It seems reasonable, therefore, to attribute greater physical significance to Feldstein's[4] auroral zone. Although its exact polar distance is representative only for periods of high solar activity (cf. Sheret and Thomas[57]), the configuration is probably valid for all levels of solar activity.

The greatest difference between Vestine's[3] and Feldstein's[4] northern auroral zones is in

the region of Hudson Bay, in North America, where Feldstein's curve reaches about 4° farther south than Vestine's.

The first attempt to investigate the auroral statistics in the southern hemisphere was made by Boller in 1898[6]. White and Geddes[7] estimated the location of the zone on the basis of scattered visual observations of aurora australis, and Vestine and Snyder[8] defined a zone based on geomagnetic observations in the Antarctic. These two zones are shown in Fig. 2 together with Feldstein's[4] I.G.Y. zone. The I.G.Y. material used was taken at 20 stations on the Antarctic continent and at two stations on islands situated south of Australia. The technique employed in the evaluation was similar to that used in the Arctic.

The approximate locations of the I.G.Y. stations are shown in Fig. 4. Between geomagnetic longitudes 120 and 230° E the stations are all situated far inside the zone.

As respects the northern hemisphere, it seems reasonable to assume that the I.G.Y. southern zone has the greatest physical significance of the three.

CALCULATIONS OF PROJECTIONS OF EQUATOR PLANE CIRCLES

Hultqvist[1] traced the geomagnetic field lines from the geomagnetic equator plane outside the earth to their intersections with the earth's surface. He used a perturbation method consisting of integration along the dipole line of the deflection from a dipole field line of the higher approximation line due to the effect of the second to the fifth terms in the spherical harmonic development of the earth's magnetic field. The calculations yielded deflection vectors for a large number of dipole lines. These gave the deflection for the point of intersection of the higher approximation field line — coinciding with the pertinent dipole line in the equatorial plane well outside the earth's surface—and the earth's surface from the corresponding point of intersection of the dipole line itself.

The calculations were carried out for 36 equidistant longitudes for each of the following geomagnetic colatitudes: 0, 2, 4, 6, 8, 10, 13, 16, 19, 22, 25, 30, 40, 50 and 60 degrees.

The results of Vestine *et al's*[9] harmonic analysis of the earth's magnetic field for epoch 1945—the most recent epoch for which data were then available—were used in the computation. Thus the position of the northern magnetic axis pole was assumed to be 78·6° N, 70·0° W. The field intensity on the earth's surface at the magnetic equator was put at 0·3119 c.g.s.

The 36 points of intersection of the higher approximation field lines crossing the geomagnetic equatorial plane along one circle give, when connected with a smooth curve, the projection of the corresponding equatorial plane circle. Fourteen such projections were calculated as described above, and from them any other high or medium latitude projection may be obtained by interpolation.

The projections are ovals with the longest diameter approximately in the plane containing the 170° and 350° geomagnetic meridians. In the northern hemisphere the ovals converge, with decreasing diameter, exactly towards the point of intersection of Bartel's eccentric dipole with the earth's surface, at which point the geomagnetic axis pole is displaced by the higher terms of the harmonic development. In the southern hemisphere the convergence point is somewhat displaced from the point of intersection of the eccentric dipole with the earth's surface.

Two circle projections are shown in Fig. 3 (northern hemisphere) and Fig. 4 (southern hemisphere). They correspond to the two circles in the geomagnetic equatorial plane having radii of 7·13 and 5·60 earth radii, for which the projections along the dipole lines are the geomagnetic colatitude circles 22° and 25° respectively.

The deflection vector tables of Hultqvist[1] also facilitate estimation of the geomagnetically conjugated point of any point at medium or high latitude by simple interpolation.

COMPARISON OF THE OBSERVATIONAL AURORAL ZONE WITH THE PROJECTIONS OF CIRCLES

(a) *Northern hemisphere*

Hultqvist[1] pointed out that the experimental northern auroral zone of Vestine[3] accorded

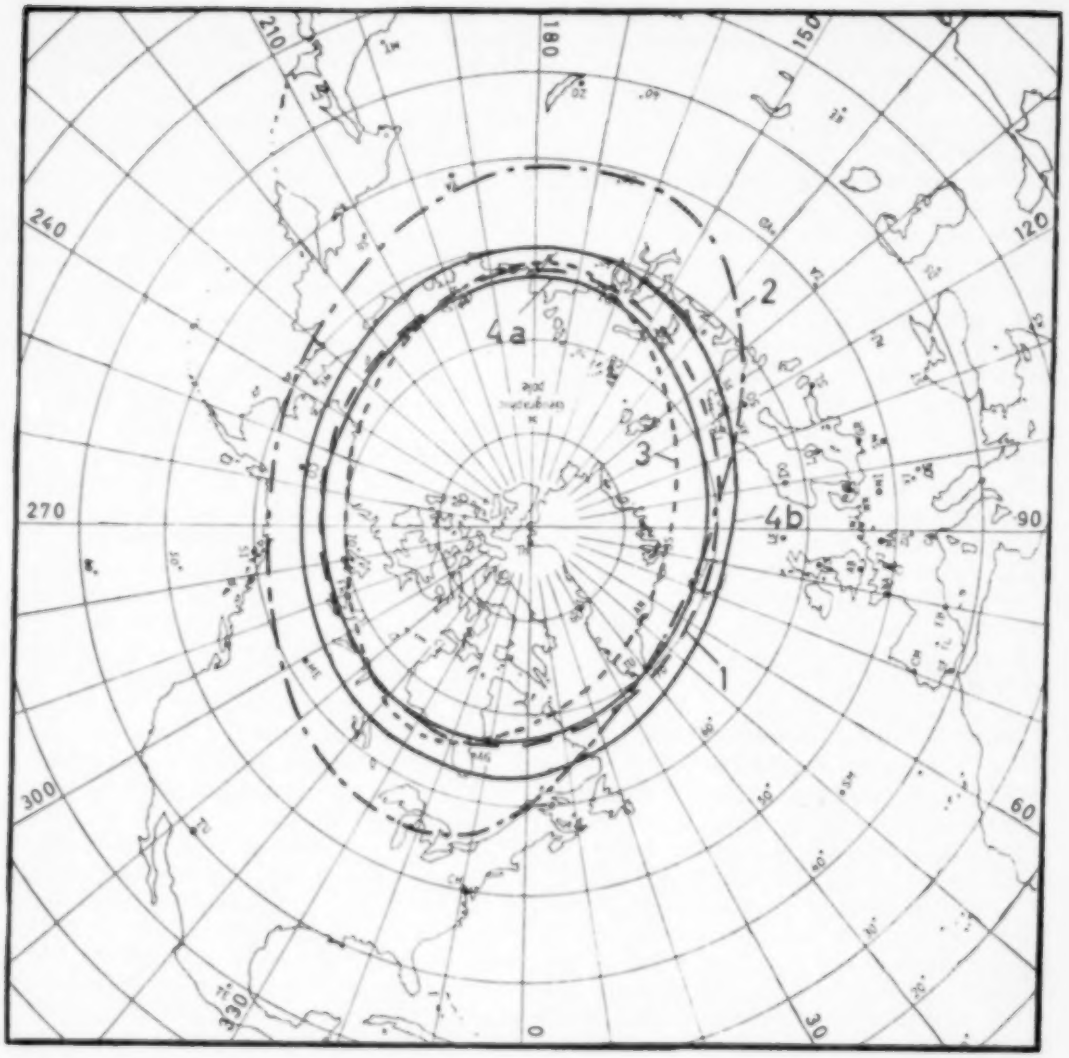


Fig. 3. Comparison of Feldstein's[4] northern observational zone (curve 1) with a number of "theoretical" auroral zones: 2 is according to Gartlein and Sprague[12], 3 to Quenby and Webber[13], and 4a and b are projections of two circles in the geomagnetic equatorial plane (corresponding to colatitudes 22 and 25 degrees respectively; Hultqvist[11]).

well with the circle projections except over the North American continent, where the calculated curve extended about 4° farther south than the observational one.

Some preliminary isochasms prepared by Gartlein[10] on the basis of visual observations in North America during the I.G.Y. agreed well with the computed circle projections (Hultqvist[11]). It seemed probable, therefore, that Vestine's auroral zone required a correction for the Hudson Bay area.

These preliminary results are verified by the

I.G.Y. data as presented by Feldstein[4]. From Fig. 3 it is evident that the configurational agreement between the I.G.Y. northern auroral zone and the projections of circles is excellent. The maximal difference in latitude between Feldstein's zone and the projection which conforms best to it amounts to about one degree.

Since the network of observation stations in and near the auroral zone was quite dense during I.G.Y., it seems probable that the error in location of the experimentally determined auroral zone is only of the order of 1° in latitude. The geomagnetic data used in the calculations are for epoch 1945, which may account for minor inconsistencies with the observations of 1957-58. It is probable, however, that this error is less than one degree of latitude (*cf.* Hultqvist[1]).

Since all the I.G.Y. observational material was obtained over a very short period of time, errors introduced by differences in solar activity for different parts of the data (*cf.* Sheret and Thomas[5]) are of no consequence.

It is evident that the I.G.Y. auroral zone is, within the limits of probable error, in full configurational accord with the calculated projections of circles in the equatorial plane.

Figure 2 includes, for purposes of comparison, curves showing two other types of more or less theoretical auroral zones: the zone suggested by Gartlein and Sprague[12] which is almost identical with the isocline 76°, and Quenby and Webber's[13] curve, which was based on calculations of vertical cut-off rigidities for cosmic ray particles in the earth's magnetic field, and took into account both the dipole and the non-dipole parts of the internal field. Neither of these two curves conforms as closely to observed data as do the circle projections.

Vestine and Sibley[14] evaluated curves nearly identical with the circle projections discussed above by computing lines of constant integral invariant, $I = \int v_{\parallel} dl / v$, on the surfaces of constant magnetic field strength, where $v_{\parallel}$ is the component of velocity along the magnetic field line and v is the total velocity. They obtained, however, in the equatorial plane, a projection which is not exactly circular but somewhat oval.

The very good agreement observed indicates that the precession of the geomagnetic axis around the geographic one has no major effects on the statistical parameter discussed.

(b) *Southern hemisphere*

Feldstein's auroral zone in the southern hemisphere is shown in Fig. 4 together with the projections of the circles of radii 5·60 and 7·13 earth radii.

The experimental curve differs quite significantly from the computed ones. While it follows the course of the latter in the geomagnetic longitude range of 320° E—90° E, it is much nearer to the Antarctic continent for the rest of the longitude circle, where it is located over the sea.

The important question is whether the discrepancy between the I.G.Y. southern auroral zone and the projections of circles is a real one or attributable merely to inaccuracy of the observational data and/or of the computed curves.

The errors of data are certainly greater for the Antarctic than for the Arctic. As mentioned earlier, the computations of circle projections were based on Vestine *et al's*[9] harmonic analysis of the geomagnetic field for epoch 1945, when the southern hemisphere data were very sparse. Although this fact may partially account for the discrepancy, it seems unlikely that the 6 latitude degrees difference over the sea is primarily attributable to errors in the harmonic analysis of the geomagnetic field.

All observational stations in the discrepant region are situated on or close to the coast of the Antarctic continent far inside the auroral zone. This means in effect that the aurorae occurring in the auroral zone are visible from these stations at low elevation, a circumstance which implies considerable error in the observational results.

Indeed, since publication of Feldstein's paper, Bond and Jacka[15] have communicated a report in which they note that the latitude of the auroral zone in the region of Macquarie Island south of Australia (point 15 on Fig. 4), is only one or two degrees more than that of the station. A small part of their auroral zone—

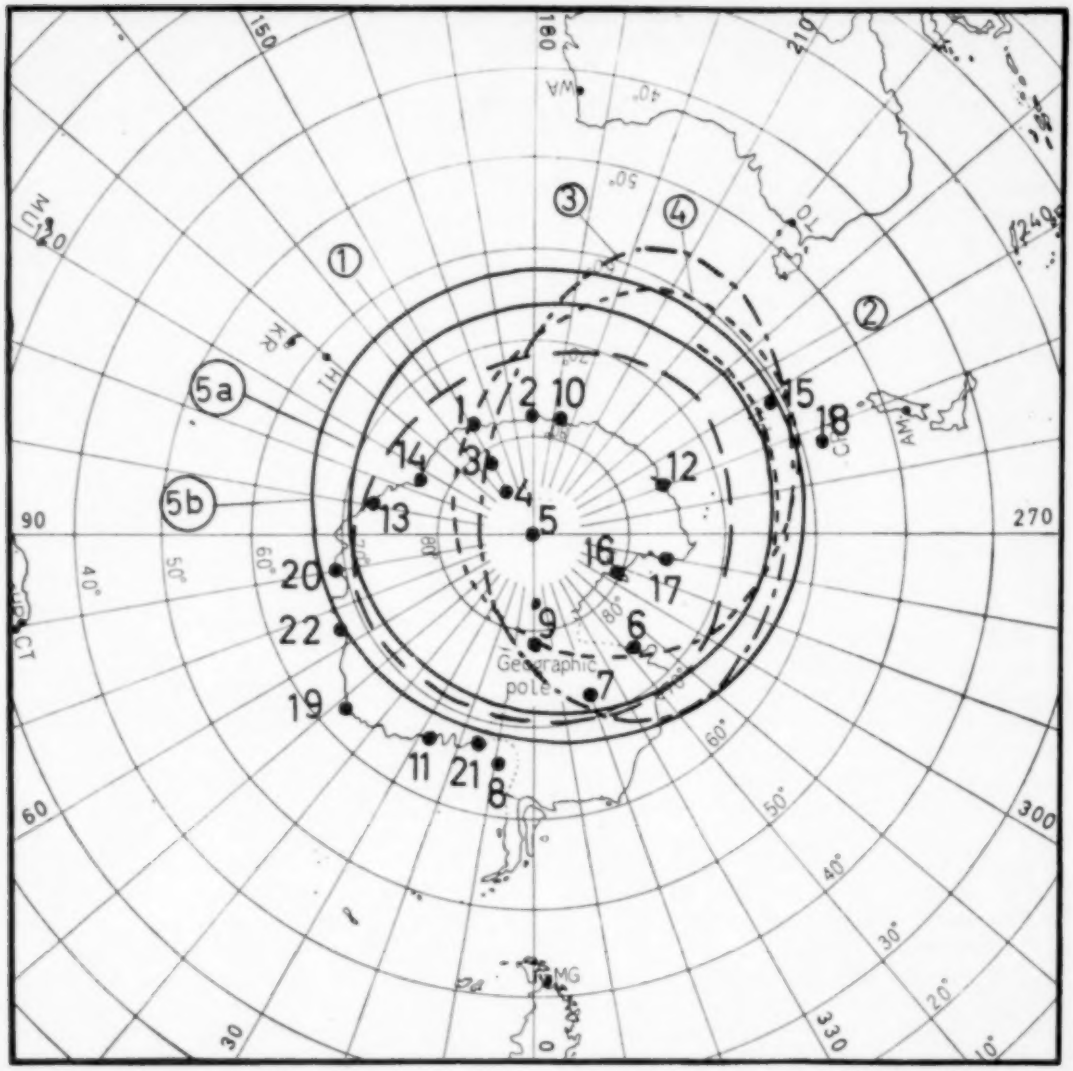


Fig. 4. Comparison of Feldstein's[4] southern observational zone (curve 1) and that of Bond and Jacka[15] (curve 2) with different "theoretical" auroral zones: 3 is according to Gartlein and Sprague[12], 4 to Quenby and Webber[13], and 5a and b are projections of two circles in the geomagnetic equatorial plane (corresponding to colatitudes 22 and 25 degrees, respectively; Hultqvist[1]). The numbered points show the approximate locations of the stations at which the IGY observational material—the basis of Feldstein's auroral zone—was collected. The numbering is that of Feldstein[4].

the one near Macquarie Island—is shown in Fig. 4 as curve No. 2. Taking this and the abovementioned considerations into account, it would appear that the auroral zone is fairly consistent with the computed curves, with which it shows agreement over the Antarctic continent. This seems so much the more plausible in view of the close accord between observed auroral zone and circle projections for the northern hemisphere. Indeed it would be

surprising if the two hemispheres showed entirely different degrees of agreement between the two types of curves.

Since the unreliability of the geomagnetic data for the southern hemisphere, and especially for the Antarctic, in epoch 1945 may play an important—though not predominant—role, a re-evaluation of the circle projections for epoch 1955 is planned.

Figure 4 contains two other theoretical auroral zones corresponding to the northern hemisphere zones of Gartlein and Sprague[12] and of Quenby and Webber[13]. The lack of agreement between observations and the two postulated zones is considerable even over the Antarctic Continent, where the observational data are certainly good.

OTHER EVIDENCE OF CIRCULAR SYMMETRY IN THE GEOMAGNETIC EQUATORIAL PLANE

Hitherto it is only the observed in relation to the computed form and location of auroral zones that have been discussed. There are, however, other statistical parameters which, according to the considerations in the Introduction, should show circular symmetry. One is the average southern extent of the aurora at a definite degree of magnetic disturbance. Such curves have been evaluated by Gartlein *et al*[17]. They can be studied in Fig 5. On close examination the curves are seen to coincide almost exactly with circle projections.

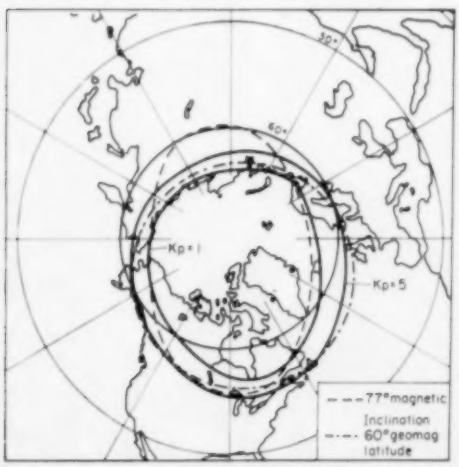


Fig. 5. Southern extent of aurora borealis for two different levels of geomagnetic activity (after Gartlein, Gartlein and Sprague[17]).

Lassen[18] has reported that morning aurorae have been observed at zenith only from places situated along the circle projection passing through Godhavn.

CONCLUSION

In the author's opinion the data discussed confirm that the northern and southern auroral zones are geomagnetically conjugated curves and are projections of circles in the geomagnetic equatorial plane on to the earth's surface along the real geomagnetic field lines.

In the southern hemisphere more accurate data are still required for those parts of the zone which lie over the ocean.

Acknowledgment—The author is indebted to the staff member of the Observatory, Mr. L. Liszka, for translation of papers published in the Russian language, to Mrs. G. Hultqvist for drawing the figures and to Mrs. G. Wirtala for secretarial help.

REFERENCES

1. B. Hultqvist, *Ark. Geofys.*, **3**, 63 (1958).
2. H. Fritz, *Das Polarlicht*, Brockhaus, Leipzig (1881).
3. E. H. Vestine, *Terr. Magn. Atmos. Elect.*, **49**, 77 (1944).
4. Y. I. Feldstein, *Geographical Distribution of Aurora and Azimuth of Auroral Arcs. Investigations of the Aurora*, The Academy of Sciences of the U.S.S.R., Moscow, **No. 4**, 61 (1960).
5. M. A. Sheret and G. M. Thomas, *Nature, Lond.*, **189**, 826 (1961).
6. W. Boller, *Beitr. Geophys.* **3**, 56, 550 (1898).
7. F. G. White and M. Geddes, *Terr. Magn. Atmos. Elect.*, **44**, 367 (1939).
8. E. H. Vestine and E. J. Snyder, *Terr. Magn. Atmos. Elect.*, **50**, 105 (1945).
9. E. H. Vestine, I. Lange, L. Laport and W. E. Scott, *The Geomagnetic Field, its Description and Analysis*, Carnegie Inst. of Washington, Publication 580, Washington D.C. (1947).
10. C. W. Gartlein, *U.S. Visual Observations, News Letter No. 18*, Ithaca, New York (1959).
11. B. Hultqvist, *Nature, Lond.*, **183**, 1478 (1959).
12. C. W. Gartlein and G. Sprague, *Auroral Occurrence*, I.G.Y. W.D.C. A., U.S. National Academy of Sciences, I.G.Y. General Report No. 12, 68-73 (1960).

13. J. J. Quenby and W. R. Webber, *Phil. Mag.*, **4**, 90 (1959).
14. E. H. Vestine and W. L. Sibley, *J. Geophys. Res.*, **65**, 1967 (1960).
15. F. R. Bond and F. Jacka, *Aust. J. Phys.*, **13**, 610 (1960).
16. K. Lassen, *Nature, Lond.*, **184**, 1375 (1959).
17. C. W. Gartlein, A. E. Gartlein and G. Sprague, The Aurora and the Local Magnetic Field, I.G.Y. W.D.C.A., U.S. National Academy of Sciences, *I.G.Y. General Report No.* 12, 57-67 (1960).

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 151–156. Printed in Great Britain.

THE ABSORPTION OF RADIO WAVES IN THE IONOSPHERE WITH RESPECT TO GEOMAGNETIC CURRENT CONTROL AND OTHER PROPOSED MODELS—A CRITICAL SURVEY

J. D. WHITEHEAD

Radio Research Board Laboratory, School of Electrical Engineering, University of Sydney, Sydney, N.S.W.

(*Received* 31 *August* 1961)

Abstract—It is shown that the arguments advanced by Shaw[1] to demonstrate that the absorption of radio waves in the ionosphere is controlled by the currents causing geomagnetic variations are unsound. Further the method used by Bandyopadhyay[2] in deducing the non-deviative absorption leads to too high a proportion of this absorption in the total. The two *D*-regions model proposed by Rumi[3] is also unsatisfactory in several respects. In all three papers, error arises because of the neglect of the deviative absorption in *E*-region. The reason for this neglect may be because of the resemblance between the frequency variation of *E*-region deviative absorption and that of the non-deviative absorption, except in the immediate vicinity of f_0E.

1. INTRODUCTION

The absorption of radio waves reflected at vertical incidence from the ionosphere has been studied for many years and the results of a comprehensive series of measurements made at Slough have been given by Appleton and Piggott[4]. Unfortunately this paper has created the impression that non-deviative absorption forms the major part of all the absorption at frequencies more than about 10 per cent away from *E*- and *F*-region critical frequencies: in fact Appleton and Piggott make the statement that "examining the noon data in detail we find that about 80 per cent of days on which accurate multifrequency measurements have been made are consistent with the example in Fig. 4," their Fig. 4 illustrating the statement made above. Whitehead[5] made a special study of the departure of the Slough measurements from the form suggested by Appleton and Piggott[4], by fitting the noon values of the absorption, $-\log\rho$, at 2 Mc/s and 4 Mc/s to the expression

$$-\log\rho = C + B(f+f_L)^{-2} \qquad (1)$$

where

f = working frequency

f_L = longitudinal gyro-frequency.

These two frequencies were chosen because only at midsummer in the sunspot maximum years did either approach to within 10 per cent of the *E*-region critical frequency. The term C represents the departure from the frequency law for non-deviate absorption which is proportional to $(f+f_L)^{-2}$.

He found that C, which was always positive when monthly means were considered, had a maximum value of about 2·5 nepers in summer throughout the sunspot cycle, and a minimum value of about 0·5 nepers in winter. In summer particularly this represents a major departure from the non-deviative law.

A more detailed analysis using Slough absorption measurements on 2, 2·4, 3·2, 4 and 4·8 Mc/s (Whitehead[6]) showed that about 50 per cent of the absorption for frequencies less than 0·7 f_0E is due to deviative absorption in a Chapman[7, 8] *E*-region. The reason for the failure of Appleton and Piggott[4] to appreciate this probably lies in the rather close resemblance of the frequency variation of the absorption for such a region to that for non-deviative absorption. In Fig. 1, $(-\log\rho)^{-\frac{1}{2}}$ is plotted as a function of frequency for a Chapman region, the dotted line showing the variation for non-deviative absorption. (Whitehead[9] showed that

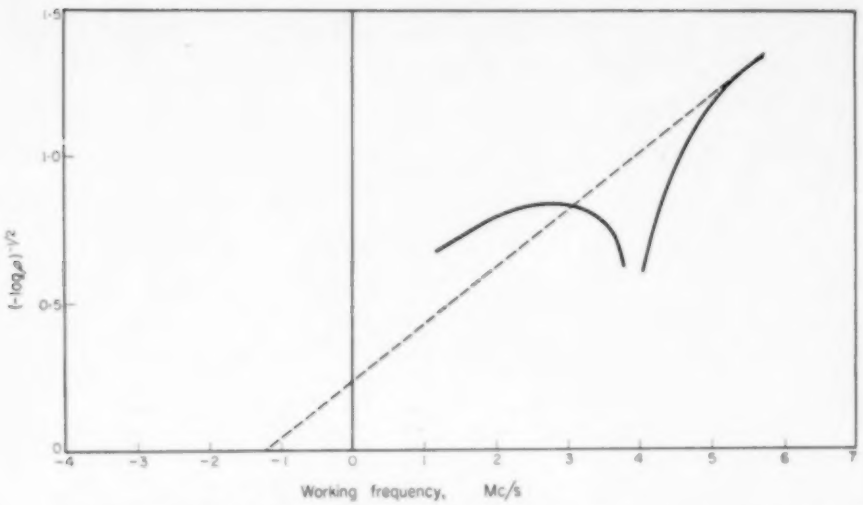


Fig. 1. Plot of $(-\log\rho)^{-\frac{1}{2}}$: working frequency for a Chapman region. The dotted lines illustrate that for non-deviative absorption.

the absorption of a wave reflected from a Chapman region occurs very close to the reflection level and must therefore be regarded entirely as deviative absorption.) It is seen that absorption is not unlike that for non-deviative. Indeed the higher value of $(-\log\rho)^{-\frac{1}{2}}$ for the Chapman region absorption than for the non-deviative at low frequencies could account for some of the observed increase below $f=1{\cdot}7$ Mc/s in Appleton and Piggott's[4] Fig. 4 (also reproduced by Rumi[3] as Fig. 3, and Bandyopadhyay[2] as Fig. 6). The jump in $(-\log\rho)^{-\frac{1}{2}}$ at this frequency may be due to reflection from a sharp thin layer within E-region (Whitehead[10]), an assertion denied by Appleton and Piggott[4], or due to errors which for a single measurement may amount to 10 per cent. It could not be due to the influence of the extraordinary wave as suggested by Rumi[3] as this would not change $(-\log\rho)^{-\frac{1}{2}}$ by more than 0·014 (as against the observed change of 0·1) nor would the influence change rapidly with frequency.

It is now possible to see how major errors could arise in the work of Shaw[1], Bandyopadhyay[2] and Rumi[3] all of whom assumed (by statement or implication) that the deviative absorption could be neglected at frequencies less than about 0·9 f_0E. The main criticism in this paper must be directed at Shaw[1], partly because of the importance of his hypothesis that the non-deviative absorption is controlled by the dynamo current system. His argument will therefore be examined in detail in the next section, and those of Bandyopadhyay[2] and Rumi[3] in succeeding sections.

2. ON THE DYNAMO CURRENT HYPOTHESIS

It is necessary first to correct Shaw's[1] apparent misinterpretation of previous work. Whitehead's[5] results showed quite clearly that the non-deviative frequency law is most closely followed in winter when C is a minimum, in contrast to Shaw's statement that this is the season when it does *not* hold. Further, the quantity C in (1), whatever else it may represent, does *not* include the seasonal variation in the non-deviative absorption, as stated.

Shaw[1] uses the parameter $f_{\min}:f_0E$ to describe the absorption at various stations. It is far from clear exactly what this represents. Whilst $f_{\min}$ measured at one station may be of some use in tracing variations in absorption (though Yasuda and Sagiuchi[11] found that

f_{min} does not even describe these very well), the change from station to station depends on many factors other than absorption.

Shaw's statement of what $f_{min}:f_0E$ represents is somewhat ambiguous, but he seems to argue that the ratio gives the non-deviative absorption in E-region, the non-deviative absorption in D-region being normalized by using the ratio rather than f_{min} itself. Now even if Shaw's assumption is true that the peak electron densities in the D- and E-regions follow the same $\cos^{2n}\chi$ law (where $n\sim 0{\cdot}3$), and if it is further assumed that $f_{min}\propto -\log\rho$ (as would be the case if the minimum detectable signal and radiated power follow power laws of frequency), there would still be a variation of $f_{min}:f_0E$ with $\cos\chi$ because the height and hence the electronic collision frequency in D-region varies with $\cos\chi$, giving in fact $-\log\rho\propto\cos^m\chi$ where $m\sim 0{\cdot}8$. Thus $f_{min}:f_0E\propto\cos^{m-n}\chi\sim\cos^{0{\cdot}5}\chi$, a law which in fact describes the way in which $f_{min}:f_0E$ varies with latitude rather well, without resort to the geomagnetic current hypothesis.

Having "normalized" f_{min} to eliminate D-region absorption, Shaw[1] then makes the statement that "deviative absorption at reflection in the E-region may be considered negligible for the f_{min} values used, because they are well below the E-layer critical frequency. Where high values of f_{min} occur, they have been limited to 0·9 (f_0E) in the analysis." The only possibility of this being so is if the E-region is extremely sharply bounded below its maximum: yet true height analyses of E-region $P'=f$ curves (Robinson[12]) consistently showed that the E-region electron density distribution has a shape not unlike a Chapman[7,6] layer and it therefore follows from the remarks in the Introduction that the absorption is not negligible for $f<0{\cdot}9\,f_0E$.

Shaw's next argument concerns the sudden jump in the ratio $f_{min}:f_0E$ at a geomagnetic latitude of 35°. This jump, I believe, is not only illusionary (when one considers the errors involved in comparing stations) but is not even required or expected on Shaw's hypothesis. It is true to the E-W S_q current reverses at this latitude, but there is no sudden change or discontinuity in it. The overall variation with latitude has already been explained without resort to the geomagnetic current influence.

The lunar tide in $f_{min}:f_0E$ does not in itself support Shaw's hypothesis, for two reasons. First the lunar tide in the geomagnetic current requires several years of data to establish (Chapman and Bartels[13]), so that it is unlikely that a lunar tide in any parameter depending on the current could be proved on the basis of two and a half year's data. Secondly, the current lunar tide at magnetically disturbed times is of the same phase as, but several times the amplitude of, that for quiet times (Chapman and Bartels[13]), whereas Shaw[1] found no lunar tide in $f_{min}:f_0E$ on disturbed days.

Shaw has kindly allowed me to use his original data in an attempt to find the origin of the lunar tide. The lunar tide for the ten quiet days of each month he found to have an amplitude of 0·022 in the value of $f_{min}:f_0E$. The quiet days however tend to occur in clusters and there is correlation both from one day to the next (short term) and because of the seasonal and sunspot variation (long term). The short term correlation has the effect of decreasing the number of independent values leading to the lunar tide making a tidal amplitude significant (at the 5 per cent level) if its amplitude exceeds 0·017 so that Shaw's tide is still significant. The effect of the long term correlation can be eliminated by using *departures* from monthly means instead of the original values in deriving the lunar tide. When this is done, the tidal amplitude falls to 0·009 and is no longer significant. The short term correlation effect still remains of course.

Thus we see that Shaw's lunar tide was probably due to an unfortunate grouping of the quiet days rather than the effect of the geomagnetic current.

The amplitude of the S_q current variations from season to season and the corresponding departure of the $f_{min}:f_0E$ ratio from the mean ("no-current") value is used by Shaw to support his hypothesis. The two are correlated of course because both have seasonal variations with a summer maximum. Again the results provide no positive evidence for a more direct link between the two.

The mechanism of absorption which Shaw suggests could not possibly lead to the observed variations in f_{min} in any case. The ionosphere is rough for most of the time in all places: the transition from rough to smooth only leads to a change in the apparent value of $-\log\rho$ of 0·26 neper (Piggott[14]) and therefore the corresponding changes in monthly means will only be a small fraction of this, whereas the latitude variation of $-\log\rho$ is the order of 2 nepers. It therefore represents a real change in absorption.

Thus we see that Shaw's evidence for geomagnetic current control of absorption does not lend support to his hypothesis and that his suggested mechanism is in any case unsatisfactory.

3. ON MODELS INFERRED FROM ABSORPTION RESULTS

Bandyopadhyay[2] calculated the deviative absorption from the noted departure of $-\log\rho$ from the non-deviative absorption law, using Appleton and Piggott's[4] result illustrated in their Fig. 4. Because of the close resemblance of the deviative absorption law to that of the non-deviative at frequencies well away from the critical frequency, Bandyopadhyay's[2] estimate of the deviative absorption is much too low, as illustrated in his Table II comparing his estimate with that for Jones'[11,13] E-region model. Because of the discrepancy, he concludes that neither the height nor shape of Jones' model is acceptable.

It is true that Jones' model is far too low and leads to too much absorption at all frequencies. However, no reasonable model is compatible with his estimate: it would require a region having a normal electron density distribution near its peak and a very sharp fall off below this. It is evident that Bandyopadhayay made an over-estimate of the proportion of non-deviative absorption.

4. ON THE TWO LAYER D-REGION MODEL

Rumi[3] uses various experimental checks to support the existence of two D-regions, in the lower of which, D_1-region, the electronic collision frequency ν_1 is a few megacycles and in the upper one, D_2-region, ν_2 is small. He first proves that instead of a straight line on a $(-\log\rho)^{-\frac{1}{2}}:f$ plot (as for normal non-deviative absorption), one obtains a point of inflexion near the frequency for which

$$2\pi(f+f_L)\sim\nu_1$$

On any assumed model of the layers, it is possible to calculate the expected $(-\log\rho)^{-\frac{1}{2}}:f$ curve. The most pronounced inflexion will occur for a model in which $\nu_2\ll f+f_L$ and the lower layer is so thin that ν_1 may be regarded as constant. A thick lower layer results in a smearing of the values of ν_1 and consequently less pronounced inflexion. For the model giving the most pronounced inflexion $(-\log\rho)^{-\frac{1}{2}}$ can be written as

$$(-\log\rho)^{-\frac{1}{2}}=\frac{K\xi}{1+P}$$

where

$$K=\left(\frac{1}{5{\cdot}3\times10^{-6}}\cdot\frac{\int_1 N_1\,dh}{\nu_1}\right)^{-\frac{1}{2}}$$

$$\xi=\Omega\sqrt{\frac{(1+\Omega^2)(1+P)}{\Omega^2+P(1+\Omega^2)}}$$

$$P=\frac{\int_2 N_2\nu_2\,dh}{\nu_1\int_1 N_1\,dh}$$

$$\Omega=2\pi(f+f_L)/\nu_1$$

and the integrals $\int_1$ and $\int_2$ are taken over the heights h of the D_1- and D_2-regions respectively. P and K are constants for the model and thus the frequency variation of $(-\log\rho)^{-\frac{1}{2}}$ depends only on that of ξ. For high frequencies $\xi=\Omega$ so that all the $\xi:\Omega$ curves tend to the same limit at high frequencies.

From the relative slopes at high and low frequencies in Fig. 2 of Rumi's[3] paper, it seems that he took $P=1$. The function ξ against Ω is plotted out in my Fig. 2 (full line) for this value of P, together with the corresponding curve given by Rumi adjusted to give the point of inflexion at the same place (dotted line). It is seen that Rumi's plot exaggerates the inflexion many times and it seems most unlikely that a small inflexion such as that on the full line of Fig. 2 could lead to the effects noted by Rumi. Plots of $\xi:\Omega$ for a large range of values of P in Fig. 3 illustrate that the inflexion is always more gradual than Rumi supposed.

Rumi[3] now postulates that the existence of the point of inflexion is checked by three experi-

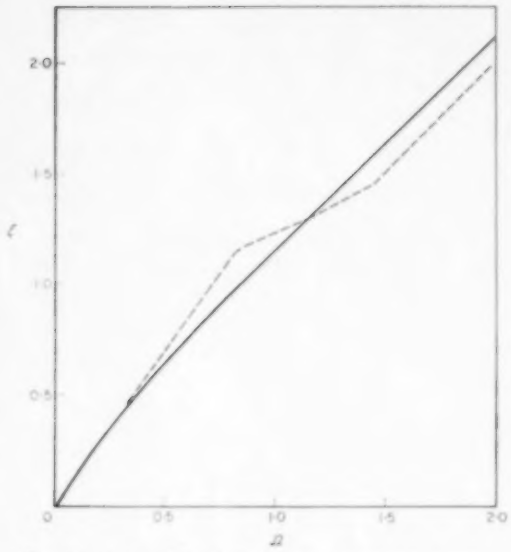


Fig. 2. $\xi : \Omega$. The dotted line shows the corresponding curve given by Rumi.

mental results. The first is the deviation from the pure non-deviative law below 1·7 Mc/s in Appleton and Piggott's[4] Fig. 3, referred to in the Introduction. As was shown there, the deviation is to be expected if a large part of the absorption was deviative. The second is that on a certain day, the 4·8 Mc/s absorption exceeded the 4 Mc/s absorption. Alternatively this may be due to extra deviative absorption in F-region which may be appreciable if $f_0F_1 \sim 4{\cdot}8$ Mc/s (Whitehead[6] estimated the F-region absorption to be ~ 1 neper even when the working frequency was not near a critical frequency). The third result is the step-like function obtained for successive quarter hour values of $f_{\min}$. This may be due to focusing effects which produce large changes in amplitude for periods of a few minutes (Whitehead[16]), or, as suggested by Yasuda and Sugiuchi[11], may be due to interference.

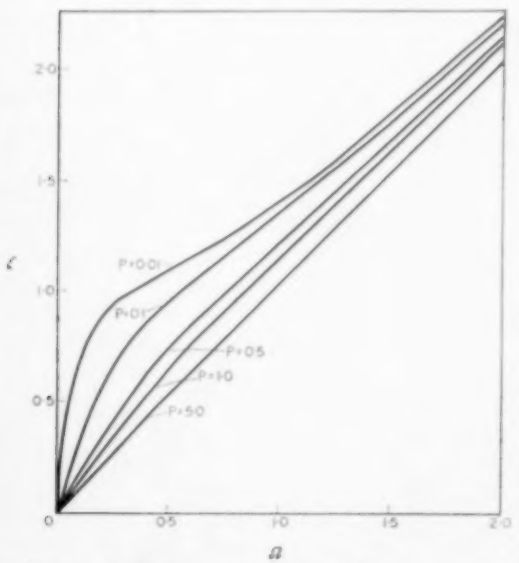


Fig. 3. $\xi : \Omega$, for various values of P.

Thus all three of the experimental results used by Rumi[3] are explicable without postulating two D-regions, and in any case it is doubtful whether the two D-region model could explain the observations.

5. CONCLUSIONS

Neglect of the deviative absorption in the E-region of the ionosphere, first of all suggested by the results of Appleton and Piggott[4], and other interpretative errors, indicate that the hypothesis suggested by Shaw[1] (geomagnetic current control of absorption), and Rumi[3] (two D-regions model) are unsound. Bandyopadhyay's[2] method of calculating the non-deviative absorption leads to a serious overestimate of this.

Acknowledgements—This work has been carried out in the School of Electrical Engineering of the University of Sydney as part of the research programme sponsored by the Radio Research Board of the Commonwealth Scientific and Industrial Research Organization.

I wish to express my appreciation to Professor W. N. Christiansen for the provision of facilities in the School of Electrical Engineering and to Mr. J. Mawdsley of the Dominion Physical Laboratories, Lower Hutt, New Zealand, for helpful discussion. I am indebted to Mr. J. E. Shaw for the loan of his original data.

REFERENCES

1. J. E. Shaw, *Planet. Space Sci.*, **2**, 1 (1959).
2. P. Bandyopadhyay, *Indian J. Phys.*, **31**, 297 (1957).
3. G. C. Rumi, *J. Geophys. Res.*, **65**, 3625 (1960).

4. E. V. Appleton and W. R. Piggott, *J. Atmos. Terr. Phys.* **5,** 141 (1954).
5. J. D. Whitehead, *J. Atmos. Terr. Phys.*, **10,** 12 (1956).
6. J. D. Whitehead, *J. Atmos. Terr. Phys.*, **16,** 283 (1959).
7. S. Chapman, *Proc. Phys. Soc. Lond.*, **43,** 26 (1931).
8. S. Chapman, *Proc. Phys. Soc. Lond.*, **43,** 483 (1931).
9. J. D. Whitehead, *J. Atmos. Terr. Phys.*, **9,** 276 (1956).
10. J. D. Whitehead *J. Amos. Terr. Phys.*, **16,** 99 (1959).
11. J. Yasuda and H. Sagiuchi, *J. Rad. Res. Labs. (Japan)*, **7,** 551 (1960).
12. B. J. Robinson, *J. Atmos. Terr. Phys.*, **18,** 215 (1960).
13. S. Chapman and J. Bartels, *Geomagnetism*, Clarendon Press, Oxford (1940).
14. W. R. Piggott, *Proc. Instn. Elect. Engrs.*, Pt. III, **100,** 61 (1953).
15. R. E. Jones, *J. Atmos. Terr. Phys.*, **6,** 1 (1955).
16. J. D. Whitehead, *J. Atmos. Terr. Phys.* **9,** 269 (1956).

SELECTED PAPERS AND ABSTRACTS TRANSLATED FROM

ARTIFICIAL EARTH SATELLITES
U.S.S.R.

(ISKUSSTVENNYE SPUTNIKI ZEMLI)

(PUBLISHED BY THE U.S.S.R. ACADEMY OF SCIENCES,
MOSCOW)

Translations of any papers abstracted here may be obtained from Pergamon Institute at a reasonable cost. Enquiries should be addressed to the Administrative Secretary, Pergamon Institute, at either Headington Hill Hall, Oxford, England, or 122 East 55th Street, New York 22, N.Y., U.S.A.

Any papers marked * have been selected for publication in this section of future issues of *Planetary and Space Science*.

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 159–164. Printed in Great Britain.

DETERMINATION OF ILLUMINATION CONDITIONS AND PERIODS OF AN ARTIFICIAL SATELLITE IN SHADOW AND IN SUNLIGHT

I. M. YATSUNSKII

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 4, p. 35, 1960.

For solving several problems connected with artificial satellites (the performance of solar batteries, thermal conditions, orbit measurements, etc.) a knowledge is required of the periods in shadow and in sunlight of a satellite, together with the possibility of observing it by optical methods. On the other hand, in order to obtain a given position of the orbital plane relative to the sun (for example, for the purpose of observing the sun itself), it is necessary in a corresponding manner to select the moment of launching into orbit.

Experience with the first artificial satellites shows that for the operational prediction of visibility conditions, for evaluating the illumination at different periods in the life of the satellite and also for selecting the moment of launching, a grapho-analytical method can be used to advantage, considering the characteristics of the satellite path with reference to the Earth's shadow.

The present paper sets out one of the possible methods for determining the conditions of illumination of a satellite and provides a brief analysis of the path of the first, second and third Soviet satellites with reference to the Earth's shadow.

1. GRAPHO - ANALYTICAL METHOD FOR ESTIMATING THE CONDITIONS OF ILLUMINATION

We will examine a single sphere with centre at the Earth's centre. From a spherical triangle with corners corresponding to the Sun, the North Pole and the satellite, for a given moment in time we can readily obtain

$$\cos z = \sin \Delta = \sin \psi \sin \delta + \cos \psi \cos \delta \cos t_{\odot}. \quad (1)$$

z is the zenith distance of the Sun at the point occupied by the satellite at the given moment ($\Delta = 90° - z$ – height of Sun); ϕ is the geocentric latitude of the satellite; δ is the declination of the Sun; $t_{\odot}$ is the hour angle of the Sun relative to the meridian of the satellite The angle $t_{\odot}$ can be related to the longitude of the satellite L_c and the time t by the formula:

$$t_{\odot} = t + 9^h + L_c - \eta \quad (2)$$

where t is the Moscow time, corresponding to the position of the satellite on a meridian with longitude L_c; η is the time "equation," i.e. the difference between mean and true solar time.

The first part of the problem is solved by equations (1) and (2), the zenith distance of the Sun being determined from the given coordinates of the satellite and the time. Obviously the solution for any point on the Earth's surface does not differ in any way from that set out. A disadvantage from the standpoint of investigating satellite behaviour is the lack of satellite-path parameters, which does not allow long-term prediction, in advance, of the condition of illumination, without preliminary and fairly accurate calculation of the entire orbit. However, as can be shown, determination of the conditions of illumination requires only a knowledge of the orientation of the orbital plane in space, that is to say, the solution of a much simpler problem.

We will write

$$t_{\odot} = \gamma + \lambda \quad (3)$$

where

$$\gamma = \gamma_0 - (\eta - \eta_0) + \int_{t_0}^{t} \omega_1(\tau)\, d\tau - \omega_2 (t - t_0) \quad (4)$$

$$\sin \lambda = \cot i \tan \phi \quad (5)$$

In equations (3-5) the angle γ is the difference

in the longitudes of the ascending node of orbit (L_Ω) and the solar meridian ($L_\odot$) at the moment in time t, that is to say, the hour angle of the Sun relative to the node of orbit ($\gamma = L_\Omega - L_\odot$), λ is the difference in longitudes of the satellite (L_c) and the ascending node ($\lambda = L_c - L_\Omega$); ω_1 is the angular speed of rotation of the node of orbit; ω_2 is the angular speed of rotation of the mean equatorial Sun, t is the time lapse in days, i is the orbital inclination; and t_0, γ_0, η_0 are initial values of the quantities t, γ, η. According to (2) and (3) we have

$$\gamma_0 = t_0 + 9^h + L_c^{(0)} - \eta_0 - \lambda(\phi_0) \qquad (6)$$

The speed of rotation of the node is determined with sufficient accuracy from the formula (considering $t_0 = 0$)

$$\omega_1 = \omega_1^{(0)}(1 + kt) \qquad (7)$$

whence for an initial speed of precession of the node we have

$$\omega_1^{(0)} = \frac{2\pi\varepsilon\cos i}{\mu T_0 p_0^2}$$

where ε and μ are constants of the Earth connected by the relationship $\varepsilon\mu^{-1} = \bar{a}^2(\alpha - m/2)$; in this case $\bar{a} = 6378{\cdot}1$ km is the radius of the equator of the Earth, α is the Earth's oblateness, m is the ratio of the centrifugal force to the force of gravity at the equator

$$\alpha - m/2 = 0{\cdot}001633.$$

Introducing the coefficient k into formula (7) we determine the relationship

$$k = \frac{1}{T}\left(2\frac{p'}{p} + \frac{T'}{T}\right)$$

where the quantities p' and T' are the reductions ("gradients") respectively of the orbital parameter p and the circling time T after one revolution of the satellite around the Earth.

With the passage of time, p' and T' on average increase monotonously (approximately linearly), if the interval of time t is sufficiently large. The fluctuations observed in these values, associated, according to some data, with variations in solar activity (sun spots) and with the rotation of the Sun around its own axis, are of relatively small amplitude (10-15 per cent) and cannot substantially influence the value of the correction term kt in equation (7). Similarly, larger periodic fluctuations, associated with change in the latitude of the orbit perigee, will be averaged out when the interval of time t is much greater than the period (T_ϕ) of these fluctuations. If both intervals of time considered are of the same order or $t < T_\phi$, then although a monotonous increase of p' and T' may not be observed, the absolute variations in these quantities in the given case are so small that, in calculating the correction term kt, it is appropriate as before to consider them as linear functions of time.

Hence we take

$$p' = p'_0(1 + \alpha_1 t)$$
$$T' = T'_0(1 + \alpha_2 t)$$

where p'_0, T'_0 are the initial "gradients" of the orbital parameter and circling time; α_1, α_2 are constant coefficients, which can be considered approximately as equal to

$$\alpha_1 \approx \frac{1}{2p'_0}\frac{dp'}{dt}, \qquad \alpha_2 \approx \frac{1}{2T'_0}\frac{dT'}{dt}.$$

In practice, for known values of the parameter p and period T in some initial interval of time of flight t_1 ($t_1 < t$), for values p', T' and α_1, α_2 mean values can be taken in this interval.

Substituting ω_1 in accordance with (7) in expression (4) and integrating (with $t_0 = 0$ and $p = p_0$, $T = T_0$), we obtain for the angle γ

$$\gamma = \gamma_0 + \omega_1^{(0)} t\left(1 + \frac{k_0}{2}t + bt^2\right) - \omega_2 t - (\eta - \eta_0) \qquad (4')$$

where

$$\frac{k_0}{2} = \frac{1}{T_0}\left(\frac{p'_0}{p_0} + \frac{T'_0}{2T_0}\right)$$
$$b = \frac{2}{3T_0}\left(\frac{p'_0}{p_0}\alpha_1 + \frac{T'_0}{2T_0}\alpha_2\right)$$

Hence, knowing, at some initial moment of time, the satellite co-ordinates $L_c^{(0)}$ and $\phi^{(0)}$ the initial values of the orbital elements T_0, p_0 and also the mean "gradients" p', T' and their variations dp'/dt, dT'/dt, it is possible for a known orbital inclination, from equations (1-5)

to calculate the zenith distance of the Sun in advance for any given moment of time with a given limitation in the value of the terms $(k_0/2)t$ and bt^2 (there must be $(k_0/2)t<0{\cdot}2$, $bt^2<0{\cdot}2$). For satellites near to the Earth it is in practice possible to predict the position of the orbital plane at one-third of the satellite life, calculating from an arbitrary origin.

In addition to the position of the orbital plane relative to the Sun, the conditions of illumination depend on the height of flight of the satellite in the regions of entry and departure from shadow. The height of flight governs the dip of the horizon for the satellite (ε) which is calculated from the formula

$$\cos\varepsilon=\frac{R}{r}=\frac{R}{p}[1+e\cos(u-\omega)] \tag{8}$$

where r is the radius-vector of the satellite, R is the mean radius of the Earth, p, e, ω are respectively the parameter, eccentricity and argument of perigee of the elliptical orbit,

$$\sin u=\frac{\sin\psi}{\sin i} \tag{9}$$

The zenith distance of the Sun at the moment when the satellite crosses the boundary of the Earth's shadow will be equal to

$$z_{\odot}^{\max}=90^{\circ}+\varepsilon \tag{10}$$

Equating $z_{\odot}^{\max}$ to the zenith distance of the Sun according to equation (1) we obtain the equation for determining the latitude of entry and departure of the satellite from the Earth's shadow

$$\cos z_{\odot}^{\max}=-\sin\varepsilon=-\sqrt{1-\left(\frac{R}{r}\right)^2}=$$

$$=\sin\psi\sin\delta+\cos\psi\cos\delta\cos[\gamma(t)+\lambda(\psi)] \tag{11}$$

Equation (11) yields also the time (hour angle for the point of passage through the boundary of the shadow at a given latitude) together with a solution to the inverse problem—determination of the initial position of the orbital plane (angle γ_0), at which the satellite at a given moment t will enter (or leave) the shadow at latitude ϕ. For the purpose of obtaining a simple method for solving the various problems in the determination of the conditions of illumination, arising in connection with launching the first Soviet satellites, special curves were constructed from equation (1) in which the hour angle of the Sun $t_{\odot}$ was plotted along the axis of abscissae, and along the ordinate axis the geocentric latitude (ϕ), the height of the Sun (Δ) serving as parameter. The curves were constructed with constant declination δ (it was sufficient to select δ in steps of 5°). In addition to the lines of equal height of the Sun, discrete positions of the orbital plane of the satellite were plotted on the graphs (dotted curves) for inclination $i=65^{\circ}$. To each of the dotted curves there corresponds a given value of hour angle of the ascending node of orbit (γ), read off on the curve along the equator from the solar meridian ($\Delta=+90^{\circ}$) to the meridian of the ascending node; for each point on the curve with latitude ϕ there is its hour angle. Fig. 1 shows an example of such a curve with $\delta=-5^{\circ}$.

For solving the main problem—determination of the height of the Sun for a given satellite and in particular the point of entry and departure from the Earth's shadow—it is necessary to start from the known co-ordinates of the satellite ($\phi_c^{(0)}$ $L^{(0)}$) at the given initial moment in time, and then to calculate the angle γ_0 from formula (6). Selecting the curve with solar declination δ corresponding to the given data, we produce on it the angle γ_0 in the form of a vertical line, which will mark the initial position of the ascending node of orbit relative to the solar meridian. From equation (4′) we carry out calculation of the subsequent positions of the node (angles γ) over a given interval of time (1-2 days) to some moment t_k (not greater than one-third the life of the satellite) and produce corresponding marks on the graphs of the type in Fig. 1 with the necessary values δ.

Each mark corresponds to some given position of the orbital plane which is found by interpolation between the supporting dotted curves. The position of the orbital plane found yields directly from the curve the height of the Sun at any latitude both on the ascending and the descending sides of the orbital loop. If the value δ on the curve is somewhat out of agree-

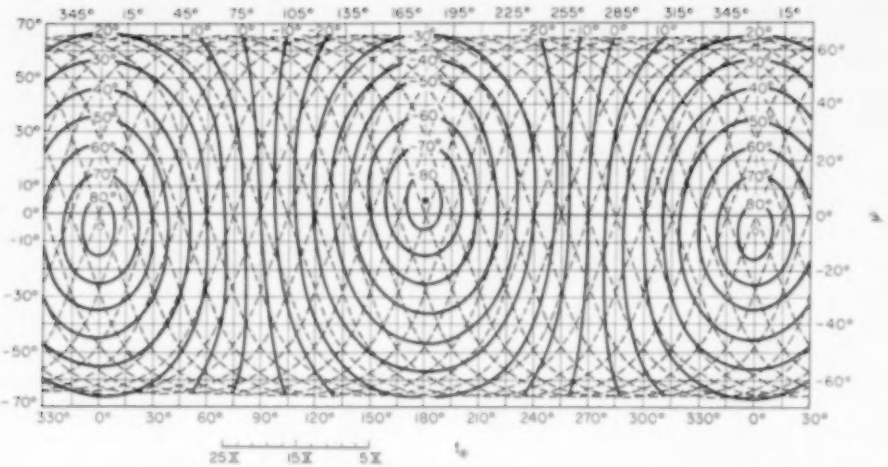


Fig. 1. Height of the Sun (declination $\delta = -5°$) as a function of the geographical latitude (φ) and the hour angle ($t_\odot$): discrete positions (every 10°) of the position of the orbital plane with inclination 65° are shown dotted.

ment with the data considered, then interpolation is carried out for δ.

Knowing the variation in the height of flight (or radius-vector r) of the satellite according to latitudes, two latitudes φ_1, φ_2 are found on the main curve for the given position of the orbital plane (entry and departure from shadow), for which the height of the Sun Δ is equal to the dip of the horizon ε (equation (8)). In practice an $\varepsilon(\varphi)$ curve can be constructed for each satellite, readily yielding the necessary comparison of Δ and ε.

The time of flight on the shadow side of the orbit is determined from the equation

$$\Delta t_{\text{shad}} = \frac{T}{2\pi}[E_2 - E_1 - e(\sin E_2 - \sin E_1)] \qquad (12)$$

where $E_2 E_1$ is the eccentric anomaly, respectively at the points of departure and entry into the Earth's shadow (determined from the values found for φ_2, φ_1). The period in sunlight will be

$$\Delta t_{\text{sun}} = T - \Delta t_{\text{shad}}$$

For carrying out optical observations on the satellite it is important to determine the "twilight" zone, i.e. the zone in which the satellite is illuminated whilst the point of observation is in the shadow. Using curves of the type shown in Fig. 1 the "twilight" zone for each position of the orbital plane is determined directly after finding the latitudes of entry and departure (φ_1, φ_2) from shadow. Sections of the orbit are marked out from points which commence with φ_1 and φ_2 to points with Δ approximately $-3°$. The region, limited on the one side approximately with a latitude φ_1 (or φ_2) and on the opposite side with the curve $\Delta \approx -3°$ and along the orbit with two equidistant curves (at a distance from the orbital plane such that the site angle β for the satellite is not less than 15-20°), will represent the zone (connected at the Earth's surface to the position of the given loop) in which optical observations are possible.

In the majority of cases (low height of flight) the "twilight" zone is determined simply as a band of latitudes confined on the one side by φ_1 (or φ_2) and on the other side by the latitude $\varphi_\Delta = -3°$ corresponding, on the curve, to the point of intersection of curve $\Delta = -3°$ with the considered projection of the orbital plane. It is understandable that, for determining the possibility of observing the satellite at the given point on a given loop, it is necessary also to check that the conditions for the site angles β ($\beta > 15$-20°) are satisfied.

The accuracy of this method is governed by

discrepancies occurring as a result of disregarding the Earth's oblateness, the finite refraction and angular diameter of the Sun. At angles of intersection of the orbital plane with the boundary of the shadow greater than 20° the corresponding discrepancy in determining the geographic latitude of entry (or departure) into the shadow does not exceed 1°. A second source of error is the lack of precise knowledge of the orbital parameters—plane orientation (γ, i) in space, the ellipse elements (p, e) and the perigee position (ω) relative to the equator, governing the height of flight at the moment of crossing the boundary of the shadow. In the case of advance prediction to one-third of the remaining life these errors may be considerable. In the majority of cases, with intersection of the shadow at angles greater than 20° errors in height of flight of 20-30 km and in the position of the orbit of 1-2° yield an error in the latitude of the point of intersection, also, not greater than 1°.

2. ANALYSIS OF THE ILLUMINATION CONDITIONS OF SOVIET SATELLITES

The conditions of illumination of the first, second and third artificial satellites in connection with launching at different times of the year and of the day on orbits with different apogee heights vary considerably from case to case. Fig. 1 shows positions of the orbital plane for the first satellite in the initial period of flight. The relative time in shadow and in sunlight for the first satellite is shown in Fig. 2. It is seen that the period varied negligibly, explained by the large angles of intersection of the orbital plane with the boundary of the shadow in the given period.

The second satellite was launched in such a way that its orbital plane was near to the

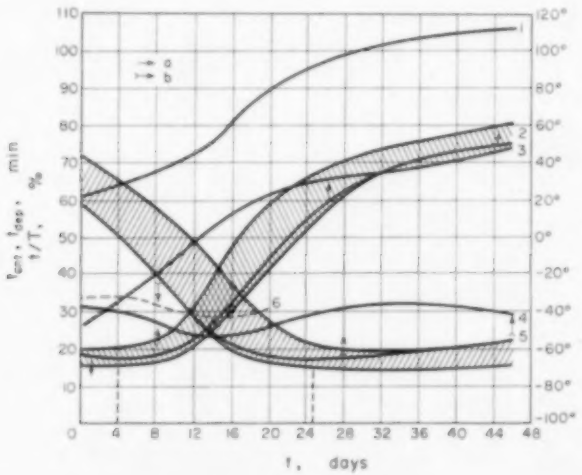


Fig. 2. Visibility conditions of the third satellite in different geographic latitudes and the period in shadow and in sunlight. 1-variation of $t_{dep.}$; 3-variation in $t_{ent.}$; 2, 5-corresponding latitudes; 4, 6-relationship $\frac{t}{T}$ for the third and first satellites. The arrows a show the moment of entry into shadow, the arrows b the moment of departure from shadow. The shaded zones represent "twilight". The vertical dotted lines are the boundaries of zones with entry and departure in opposite directions.

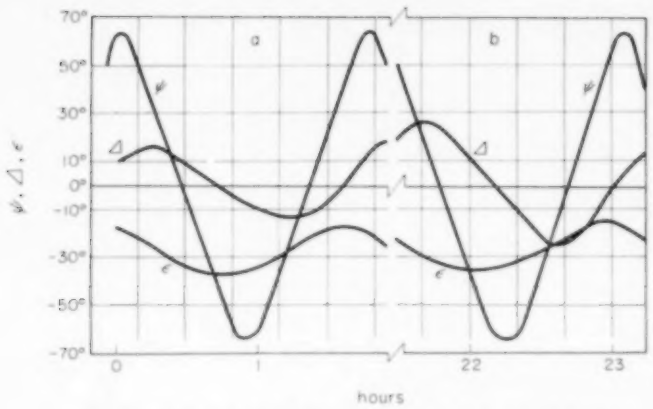


Fig. 3. Variation in height of Sun (Δ) and dip of horizon (ε) for the second satellite in the first days of flight. (*a*—3/11 1957, *b*—7/11 1957).

boundary line of the shadow. For several days the satellite did not enter the Earth's shadow. Fig. 3 shows the actual (Δ) and necessary height of the Sun (ε) for coinciding with the shadow. It can be seen that only on the fourth day after launching did the satellite begin to enter shadow. These conditions are the most favourable for the performance of the equipment intended for studying solar ultraviolet emission.

The characteristics of the path of the third satellite relative to the Earth's shadow during the first 1½ months of its life (curves of latitude-change and times of entry and departure from shadow, together with the relative period in shadow as a percentage of the circling time around the Earth) are shown in Fig. 2. Twilight zones are marked out, in which observation of the satellite was possible. Analysis of the curves shows that, as in the case of the first satellite, the period in shadow and in sunlight was fairly stable, this being connected with the large angles of intersection of the orbital plane with the shadow in the given period. Displacement of the shadow zone itself (and the twilight zones) through the latitudes was very considerable, which is explained by the relatively rapid rotation (precession) of the orbital plane.

It is observed in conclusion that the inclination of the orbital plane to the equator (65°) selected for the Soviet satellites is governed by two periods of the year during which it is possible to remain outside the shadow: the summer period from the end of April to the middle of August and winter—from the end of October to the middle of February. For great heights of flight in the region of shadow (greater than 800 km) brief periods outside the shadow are possible at any season of the year. In accordance with the circling periods of the node of orbit relative to the Sun (for the Soviet satellites approximately 3 months) existence outside the shadow will be repeated with the same or twice the frequency (in the case of great heights of flight).

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 165–172. Printed in Great Britain.

MEASUREMENTS OF COSMIC RAYS ON GEOPHYSICAL ROCKETS

Yu. G. SHAFER and A. V. YARYGIN

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 4, p. 184 (1960)

In 1958 measurements of the global intensity of cosmic rays were carried out by means of standard geiger counters and an ionization chamber installed on a geophysical rocket.

The investigations aimed at establishing the relationship of altitude to intensity up to 200 km and determining the mean specific ionization of primary emission particles. Our task also included the testing of equipment under typical overload and temperature conditions as encountered in rocket flights.

The equipment was installed in the hermetically sealed nose of the rocket, the thickness of the steel walls being not greater than 3 mm ($\sim 2{\cdot}5$ g/cm^2); the instrument axis coincided with the rocket axis. About 8 per cent of the total solid angle of the outer hemisphere was screened with a large quantity of material (> 50 g/cm^2). The solid angle, in which the equipment was closed off by the body of the rocket on the lower side of the hemisphere, was $0{\cdot}16\pi$. A diagram of the measuring equipment arrangement is shown in Fig. 1.

The particles were recorded by halogen counters. The geometrical dimensions of the counter were 18×150 mm and the thickness of the steel wall 0·1 mm. Important characteristics of this type of counter are stable performance with temperature fluctuations from $+50$ to -40°C, a prolonged service life and low operating voltage (380–400 V). Electricity supply to the counters was self-contained from converters based on semi-conductors. At the converter input a voltage of 6·5 V was applied with a current of 7·5 mA, providing an output of 400 V at 50 μ A. A gas stabilizer was installed for reducing voltage fluctuations at the converter output. The voltage, stabilized within 1–2 per cent, was independent of variations in the load current and input voltage.

The global intensity in numbers of charged particles measured by the standard counter is calculated from the relation:

$$J = \frac{N}{G} \text{ particles cm}^{-2} \text{ sec}^{-1},$$

where N is the number of particles recorded by the counter in 1 sec, G is a geometrical parameter, being a function of the angular distribution of radiation, the counter orientation, its effective length L_{ett} and effective radius r_{ett}.

On the basis that the angular distribution of primary emission on the upper hemisphere is isotropic and that the intensity is zero from the lower hemisphere, we obtain for a cylindrical counter, with arbitrary orientation in space $G = S_{ett}$ (effective area of the counter):

$$S_{ett} = \frac{\pi r_{ett}}{2}(r_{ett} + l_{ett}). \qquad (1)$$

It is known that the effective diameter of a geiger counter differs from the actual diameter by not more than 0·1 mm. The effective length of the type of counter used by us has been determined experimentally by Yu. I. Logachev and came out at (100 ± 1) mm. Writing in formula (1) the values r_{ett} and l_{ett}, we obtain $S_{ett} = 15$ cm^2. It was necessary in our experiments to take into account the "dead period" of the counter τ, that is to say the time during which the counter was unable to record.

For the counter employed by us $\tau = 2 \times 10^{-4} s$. The value τ determines the value of relative discrepancy σ in the counting rate:

$$\sigma \cong N\tau.$$

With maximum counting rate $N = 40$ particles in 1 sec, $\sigma = 0{\cdot}8$ per cent. In the further calculations we will disregard this insignificant error caused by particle counting loss.

It is very important to take into account

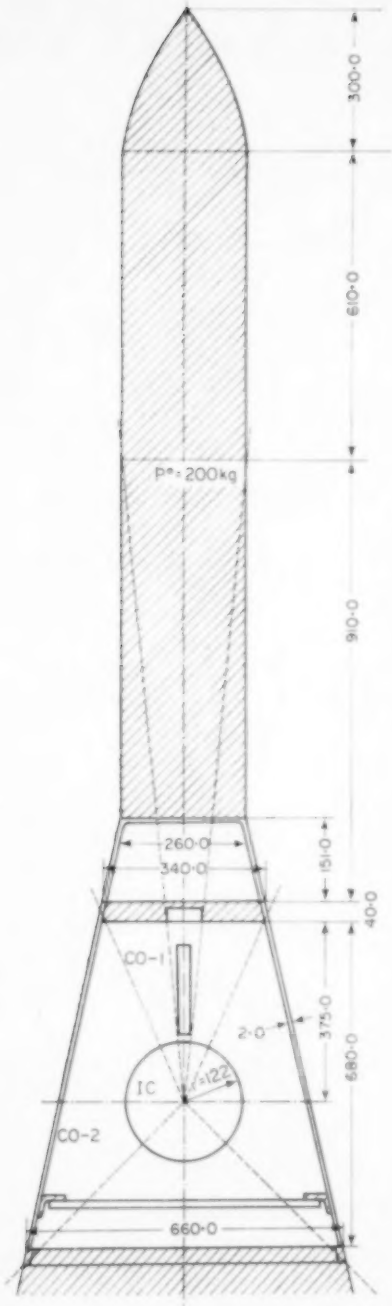


Fig. 1. Arrangement of horizontal counter (CO-2) and vertical counter (CO-1) and the ionization chamber IC in the nose of the geophysical rocket.

angular displacements of the body of the rocket during the period of free flight. Since the installation is not spherically symmetrical, variations can occur during the measurement process connected with random motion of the rocket.

Even with constant primary radiation intensity the theoretical calculation of these variations is very complex. We therefore installed in the rocket two counters, the axis of one lying along the axis of the rocket and the axis of the other perpendicular. With this counter arrangement the unaccounted angular displacement of the rocket will have a different effect on their counting rate.

The systematic difference between N_{vert} and N_{hor} may provide the possibility of determining the effect of screening of the counter with the material of the rocket and the lower accuracy limit. In addition the indications of the two counters taken from independent radio circuits and supply sources afford a check on their correct operation.

In our equipment one counter was placed vertically above the ionization chamber and the other horizontally at the side of the chamber (see Fig. 2).

The impulses from the counter arrive via an emitter-repeater at the input of the counting circuit with a counting factor 1:256. Each trigger cell of the counting circuit provides a counting rate up to 5×10^{-4} impulses in 1 sec. With this resolving capacity all the impulses arriving will be recorded with our equipment.

The final trigger of the conversion circuit controls the polarized relay (P_1) via the emitter-repeater. A change in the position of the armature of the relay occurs after storage of 128 impulses (Fig. 3). The contacts of the relay (P_1) close the appropriate potentiometer circuits of the telemeter system installed on the rocket. As the result of tripping of the final trigger on summating 128 impulses there is transmitted by radio from the rocket a given voltage level for each counter, which is recorded on a film in the receiving-station recording instrument.

Simultaneously a continuous recording of the time scale is produced on the film with an

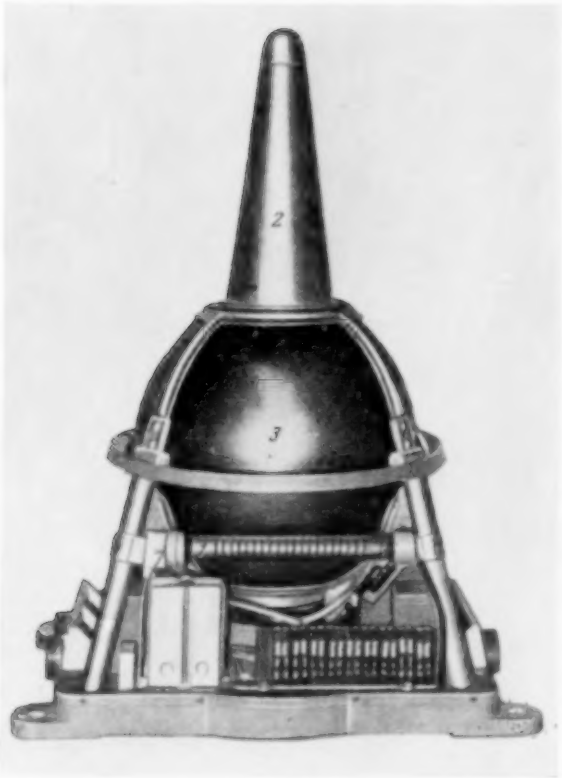


Fig. 2. Equipment for recording cosmic rays, installed on the geophysical rocket.
1—horizontal counter; 2—vertical counter (in thick-walled duralumin jacket); 3—ionization chamber.

(*To face page* 166)

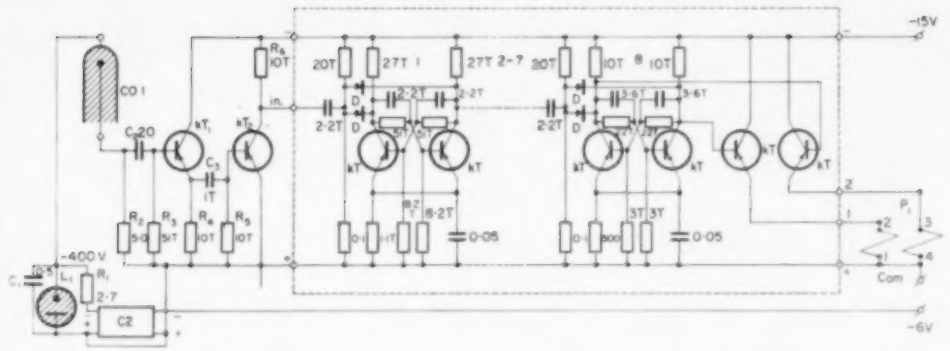


Fig. 3. Line diagram of the counting equipment.

accuracy of 0·02 sec. A specimen of this type of photographic record is shown in Fig. 4.

The number of particles recorded by the counter in 1 sec can be calculated from the relation $N=n/t$ where n is the conversion factor and t is the time measured on the film between the moments of voltage-level variation from the given counter.

In our equipment the horizontal counter at the same time performs the function of master counter. On summating 256 impulses the relay P_4 switches the voltage to the ionization chamber relay P_5 (Fig. 5). This relay functions and absorbs the charge accumulated through ionization at the collecting electrode of the chamber during the period allowing the passage of 256 particles through the master counter.

The spherical ionization chamber with a diameter of 24 cm is prepared from stainless steel 0·4 mm. thick. A chamber of this volume provides a statistical accuracy not greater than 5 per cent, governed by the fluctuations in the number of primary relativistic charged particles passing through the chamber during each second of observation.

A vibration-proof collecting electrode in the form of a hollow, thin-walled sphere with diameter 28 mm is mounted on a stem passing through a central ceramic insulator with a guard ring. The chamber was degasified, evacuated, flushed several times and filled with spectrally pure argon. Saturation curves were taken during filling after each atmosphere. The following are the ionization chamber characteristics:

(1) Volume $V=7020$ cm^3; (2) argon excess pressure $P=7$ atm; (3) resistance of the central insulator $R=1{\cdot}5\times10^{15}$ ohms; (4) conversion factor from argon to air $K=1{\cdot}4$; (5) saturation potential $U_s=250$ V; (6) working voltage $U_p=$ 380-400 V; (7) ionization current in the order of $I=10^{-11}$ a; (8) electrical capacitance $C=6{\cdot}5\times10^{-12}$ F.

For the purpose of calculating the possible statistical error occasioned by fluctuations in

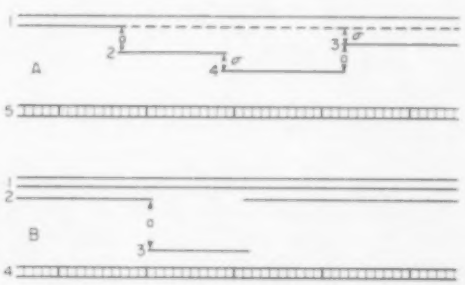


Fig. 4. Specimen photographic records of the signals provided by the equipment at the receiving station.
A—counter signals: 1—zero level; 2—signal level of the vertical counter (amplitude a); 3—signal level of the horizontal counter (amplitude b); 4— summated level of the signals from both counters (a+b); 5—time base.
B—ionization chamber signals: 1—calibration level (100%); 2—reference level; 3—ionization chamber signal (a—signal amplitude); 4—time base.

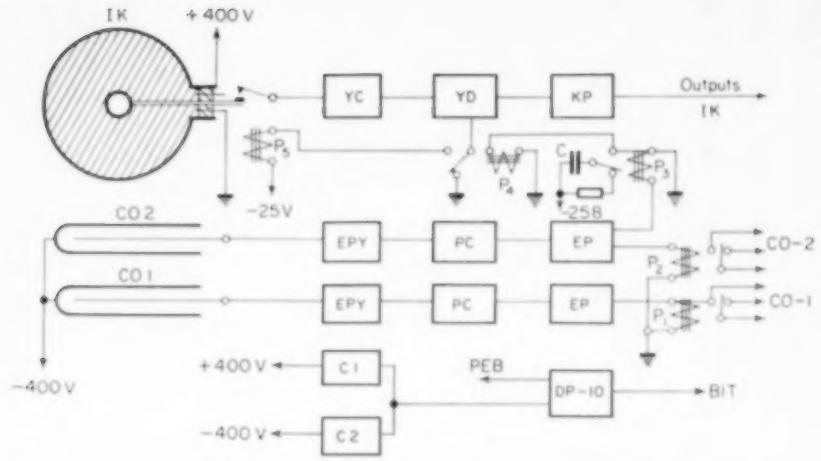


Fig. 5. Block diagram of the equipment for measuring the intensity of cosmic rays.
EPY—emitter-repeater and amplifier; *PC*—conversion circuit; *EP*—emitter-repeater; *C1*, *C2*—converters; *DP*—10—supply switch relay; *IK*—ionization chamber; *CO*-2—horizontal counter (pilot); *CO*-1—vertical counter; P_5—ionization chamber relay; *YC*—amplifier; *YD*—attenuator; *KP*—cathode repeater; P_1, P_2, P_3, P_4—polarized relays; *PEB*—electronic unit supply; *BIT*—electricity supply unit.

the number of multiple charged particles of the primary flux, the following primary emission composition was taken by us: protons—80 per cent of the total number of particles arriving from all directions, helium nuclei—18 per cent, *LiBe*, *B*-0·3 per cent, *C*, *N*, 0-1·2 per cent and other nuclei with $C \geqslant 9$—approximately 0·5 per cent.

It is known that the ionization set up by a particle is proportional to the square of its charge. The contribution of ionization of primary nuclei calculated by us to the total ionization current of the chamber will be respectively:

$$I_p = 24 \text{ per cent}; \quad I_\alpha = 21{\cdot}5 \text{ per cent};$$
$$I_{Li,Be,B} = 2{\cdot}2 \text{ per cent};$$
$$I_{C,N,O} = 23 \text{ per cent}; \quad I_z \geqslant 9 = 29 \text{ per cent}.$$

As the result of calculations the mean statistical accuracy was found to be:

$$\overline{\sigma} = \sqrt{\sum \left(\frac{\Delta I_z}{I_z}\right)^2} = 8{\cdot}3 \text{ per cent}$$

where $\frac{\Delta I_z}{I_z}$ is the relative statistical error introduced by fluctuations in the number of primary nuclei with charge Z.

The method of measuring ionization of the chamber which most justifies itself is the method where the electrical charge is taken from the collecting electrode by means of a relay. This method was first developed and applied by A. E. Chudakov[1].

In our equipment the relay of the chamber P_5 (Fig. 5) connects the collecting electrode to the input of the amplifier (*Yc*). At the load resistance of the chamber there then occurs an impulse of positive polarity with a duration of approximately 1 msec. This impulse reaches the input of the amplifier with intense feedback with an amplification factor equal approximately to 6. The signal is further attenuated by the diode attenuator *Yd*, having a time constant in the order of 100 seconds. By means of the polarized relay P_4 this impulse is chopped. In this way rectangular signals U_{out}

with a duration of approximately 1 sec reach the input of the telemetering system via the cathode repeater *KP*, being graduated in volts. These signals are then transmitted by radio and recorded on the photographic film at the ground observation station. As already shown, the work of the chamber relay is programmed by the horizontal master counter. This occurs with tipping of the armature of the relay P_3, the winding of which is connected to the output of the emitter-repeater *EP*. The charge current of the capacitor *C* switches the relay P_4. The attenuator circuit then opens and a voltage is applied to the ionization chamber relay. Since the impulse from the cathode repeater output has negative polarity the peformance of the last stage of the electronic unit of the chamber was selected in such a way as to provide a strictly constant reference voltage level with positive polarity at its output. In our experiment the reference voltage level was +5·65 V. The negative signal from the chamber was read off from this level. The electronic unit of the chamber was made up from small directly-heated low-consumption valves.

Figure 4 shows a specimen photographic recording at the ground station of chamber signals U_m, transmitted by radio from the rocket. In order to determine the value U_{out} in volts it is necessary to measure the amplitude of the rectangular impulse on a scale, determined according to the position on the film of the calibration level of the known voltage of a standard battery U_a, carried on the rocket, and of a known reference voltage level U_r, arriving from the output of the cathode repeater of the chamber channel in the intervals between the chamber signals.

In our experiment $U_a = 6{\cdot}18 \pm 0{\cdot}05$ V and $U_r = 5{\cdot}60 \pm 0{\cdot}08$ V; U_{out} is calculated from the formula $U_{out} = U_r - U_m$, where U_m is the signal measured on the photographic record.

The potential of the collecting electrode U_{in} is determined from the calibration curve $U_{in} = f(U_{out})$, being the rated characteristic of the instrument (Fig. 6).

The total ionization *I*, set up by the primary nuclei, passing through the chamber during the

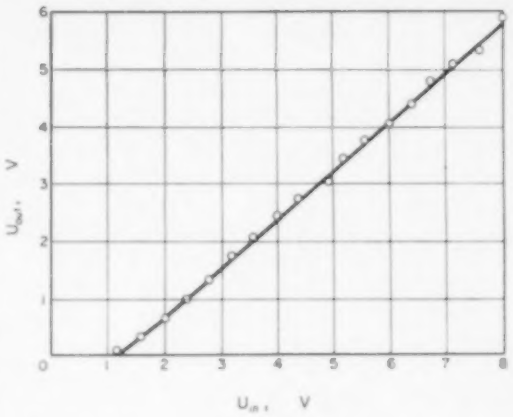


Fig. 6. Calibration curve $U_{in} = f(U_{out})$.

exposure time *t* is calculated from the formula:

$$I = \frac{U_{in}C}{Pte},$$

where *P* is the pressure in the chamber, *C* is the capacitance of the chamber, and *e* is the electron charge.

The instrument was supplied from batteries of mercuric oxide cells. The total energy consumption of the instrument was 2·1 W.

The flight equipment underwent an entire set of tests in the process of which careful checks were taken on all the circuits including the insulation, measurements were taken of the reference voltage at the cathode repeater output and the standard battery voltage, the stray resistance of the collecting electrode was checked together with the gas pressure in the chamber and the efficiency of the conversion circuits.

The counters and the chamber were calibrated by means of a gamma preparation Co^{60} and against the natural background. For the counters the natural background is equal to 70 impulses and for the chamber 0·48 V per minute. During the course of the tests the calibration curve was repeated several times with different supply voltages varying within 10 per cent from the rated.

The ionization level and counting rate were transmitted from the rocket during the tests by

radio and were recorded on the film. Identity of results of all the tests confirmed the efficient working order of the equipment.

The rocket was launched on 2nd July 1958 at 0900 hr Moscow time. During the flight the operation of the counters and chamber were checked visually on the oscillograph screen of the receiving station over a period of 264 seconds. It was found from these observations that the vertical counter operated normally throughout the whole flight. Signals from the horizontal counter and from the ionization chamber began to reach the receiving station from 213 seconds. All the equipment performed normally up to the end of the flight.

At the peak of the trajectory on reaching a height of 210 km, after the nose of the rocket had left the body, the transmission of signals by radio ceased. However, the equipment remained switched on and was successfully dropped by parachute in working condition.

As the result of careful tests carried out under field conditions on the landing site and in the laboratory we did not detect any abnormalities in the performance of the instrument.

The data obtained during the flight did not afford a decision as to whether there was a marked deviation in the intensity of cosmic rays from the mean value at the time of launching the rocket. From measurements at the Scientific Research Institute of Terrestrial Magnetism, Ionosphere and Radiowaves station in Moscow it is known that in the period from 30th June to 7th July the intensity of the hard and neutron components maintained constant values; the number of sun spots was relatively small (198 with a mean monthly value 238). Disturbance of the Earth's magnetic field and the ionosphere remained at the level of the month of June during this period. Only at the end of July were four disturbed periods detected, connected probably with some increase in solar activity.

This shows that the day was relatively calm when the rocket was launched.

One important result of our measurements was the acquisition of data relating to the

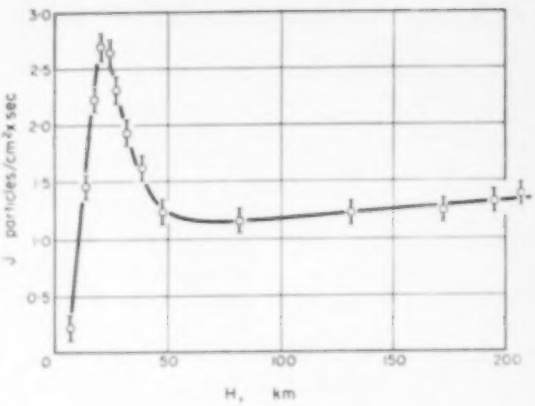


Fig. 7. Global intensity of cosmic rays as a function of altitude.

global intensity of the particle flux as a function of altitude These results are shown in Fig. 7 and in Table 1. Fig. 7 demonstrates that the counting rate at the maximum of the Pfotser curve is approximately 2·3 times greater at an altitude of 20·5 km than at 70 km. Starting from 70 km the intensity increases in monotonous fashion with altitude (approximately by 17 per cent from 70 km to 210 km).

The considerable mean square error in the region of the altitude 200 km can possibly be explained by the random variation in the position of the rocket in free flight with reduction in its velocity at the peak of the trajectory.

The intensity measured by us at the maximum of the altitude curve was 20 per cent greater than the data obtained by A. E. Chudakov[1] and 30 per cent greater than the intensity measured by A. N. Charakhch'yan in 1958 at the Simferopol latitude*. This increase is probably due to our instrument taking account, in addition, of shower particles formed by cosmic rays in the rocket material or the influence of X-ray emission.

In our experiment some part may be played by secondary particles originating in the walls of the rocket and high-power showers from the nose of the rocket. Since outside the atmos-

* From a private communication by A. N. Charakhch'yan.

phere the particle flux consists mainly of protons, the range of which is approximately 67 g/cm², then with a rocket wall thickness $d \approx 3$ mm (2·5 g/cm²) approximately 4 per cent of the particles will react with the iron nuclei. Each of these particles will produce approximately 5 new particles, of which not less than 30 per cent will be neutral. Hence the contribution of particles due to secondary particles originating in the rocket walls will be 4 per cent $\times 5 \times \frac{2}{3} \approx 13$ per cent.

The effect of showers from the nose of the rocket ($d \approx 150$ g/cm²) is associated with the reaction of high-energy particles ($\sim 10^{11}$ eV), the flux of which in primary emission is 1 per cent. Each such particle produces approximately 100 new particles. Having in mind that the solid angle which the nose of the rocket cuts out from the upper hemisphere is approximately 0·05 π, it is found that this effect to a rough estimate is equal to $1 \times 0{\cdot}05\ \pi \times 100 \approx 16$ per cent. This figure can be reduced if account is taken of the fairly small separation of the counters from the lumped mass of the rocket nose. The angle of dispersion of shower particles in our experiment was small, and therefore not every particle of the shower would be summated separately. Taking this into account it is proposed that the contribution of particles due to showers falls to 8 per cent.

These calculations provide a preliminary assessment of the upper limit of the contribution of "excess" secondary particles. It is approximately 20 per cent. Taking this into account, the total intensity beyond the range of the atmosphere, on the basis of our measurements, at an altitude of 50-70 km, is equal to

$$0{\cdot}97 \pm 0{\cdot}06 \frac{\text{particles}}{\text{cm}^2\ \text{sec}}.$$

As the result of our experiment the mean specific ionization of primary particles per 1 cm³ argon was obtained for normal pressure as $J = 260 \pm 19$ pairs of ions cm^{-3} sec^{-1}.

This value is 1·8 times greater than the mean ionization of charged relativistic particles.

The specific ionization per 1 cm path in argon with normal pressure as measured by us was $189{\cdot}1 \pm 8{\cdot}2$ cm^{-1}.

Another important result was the demonstration of the high reliability of the ionization chamber together with its electronic units and its insensitivity to the supply voltage over a range up to 20 per cent. This was confirmed by repeated checks before and after the flight on the shape of the impulses, taken from the collecting electrode of the chamber. These impulses were each time strictly identical in shape, were purely exponential, without overshoot or splitting up and were of constant duration. The same refers to the output impulse from the chamber. All the relays and relay tripping systems of the chamber, in conjunction with the master counter, operated very reliably.

Acknowledgments are made to S. N. Vernov and N. L. Grigorov for continuous guidance

Table 1. Data Relating the Total Intensity of Cosmic Rays to Altitude

H—altitude, km	J—intensity of cosmic rays, $\frac{\text{particles}}{\text{cm}^2\ \text{sec}}$	H—altitude, km	J—intensity of cosmic rays, $\frac{\text{particles}}{\text{cm}^2\ \text{sec}}$
6·18	0·21±0·01	38·18	1·63±0·10
14·15	1·46±0·09	47·73	1·25±0·07
17·64	2·21±0·13	82·00	1·17±0·06
20·50	2·69±0·16	130·82	1·25±0·07
24·00	2·63±0·16	171·70	1·27±0·12
27·00	2·30±0·12	195·10	1·34±0·12
31·75	1·94±0·12	207·40	1·40±0·16

and to A. E. Chudakov, A. I. Kuz'min and Yu. I. Logachev for valuable advice and assistance rendered.

The author expresses his thanks to Radio Engineer V. A. Belomestnyi, who rendered important practical assistance in the work.

REFERENCE

1. S. N. VERNOV and A. E. CHUDAKOV. *The study of cosmic rays by means of rockets and satellites in the U.S.S.R.*, Paper to the 5th International Assembly of the International Geophysical Year, 1958.

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 173–178. Printed in Great Britain.

INFLUENCE OF THE EARTH'S OBLATENESS ON THE ROTATIONAL MOTION OF AN ARTIFICIAL SATELLITE

V. A. SARYCHEV

Translated by R. MATTHEWS *from Iskusstvennye Sputniki Zemli* No. 6 p. 3 1961

We will determine the three right rectangular systems of co-ordinates necessary for drawing up the equations of motion (see Fig. 1).

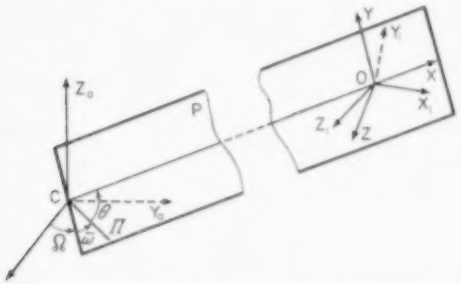


Fig. 1 Principal reference systems.
P - plane of the osculating orbit.

1. $CX_0Y_0Z_0$ is the absolute system of co-ordinates; CZ_0 is the axis of rotation of the earth; the CX_0Y_0 plane coincides with the equatorial plane.

2. $OXYZ$ is the orbital system of co-ordinates; the OX axis is directed along the radius vector connecting the mass centres of earth and satellite; OY is perpendicular to OX and lies in the plane of the osculating orbit; the OZ axis completes the selected plane n-th vector to the spatial right system of co-ordinates.

3. The axes of the $OX_1Y_1Z_1$ system of co-ordinates are directed along the main axes of the central ellipsoid of intertia of the artificial satellite; the point O is the mass centre of the satellite.

Symbols for the transition matrix elements between the main systems of co-ordinates are taken from the tables:

	x_1	y_1	z_1
x	a_{11}	a_{12}	a_{13}
y	a_{21}	a_{22}	a_{23}
z	a_{31}	a_{32}	a_{33}

	x	y	z
x_0	c_{11}	c_{12}	c_{13}
y_0	c_{21}	c_{22}	c_{23}
z	c_{31}	c_{32}	c_{33}

The direction cosines of the angles between the axes of the trihedrals $OXYZ$ and $CX_0Y_0Z_0$ can be readily expressed by the pure anomaly θ, and the osculating elements of the trajectory of the satellite mass centre [1]:

$$\left.\begin{aligned}
c_{11} &= \cos u \cos\Omega - \sin u \sin\Omega \cos i,\\
c_{12} &= -\sin u \cos\Omega - \cos u \sin\Omega \cos i,\\
c_{13} &= \sin\Omega \sin i,\\
c_{21} &= \cos u \sin\Omega + \sin u \cos\Omega \cos i,\\
c_{22} &= -\sin u \sin\Omega + \cos u \cos\Omega \cos i,\\
c_{23} &= -\cos\Omega \sin i,\\
c_{31} &= \sin u \sin i,\\
c_{32} &= \cos u \sin i,\\
c_{33} &= \cos i.
\end{aligned}\right\} \quad (1)$$

Where $u = \varpi + \theta$, ϖ is the slant range of the perigee Π from the node, u is the latitude argument, Ω is the longitude of the ascending node, i is the orbital inclination.

From Poisson's kinematic relations for direction cosines (1) expressions can be obtained for projections p_1, q_1, r_1 of the absolute angular velocity of rotation of the orbital system of co-ordinates on the OX, OY, OZ axes:

$$\left.\begin{aligned}
p_1 &= c_{31}\frac{d\Omega}{dt} + \cos u \frac{di}{dt}\\
q_1 &= c_{32}\frac{d\Omega}{dt} - \sin u \frac{di}{dt}\\
r_1 &= c_{33}\frac{d\Omega}{dt} + \frac{du}{dt}
\end{aligned}\right\} \quad (2)$$

The position of the satellite relative to the orbital system of co-ordinates is determined by means of the three independent angles ψ, θ, γ: ψ is the angle between the projection of OX_1 on the OXZ plane and the OX axis; θ is the

	x_1	y_1	z_1
x_0	$\bar{a}_{11}$	$\bar{a}_{12}$	$\bar{a}_{13}$
y_0	$\bar{a}_{21}$	$\bar{a}_{22}$	$\bar{a}_{23}$
z_0	$\bar{a}_{31}$	$\bar{a}_{32}$	$\bar{a}_{33}$

δ

angle formed by the OX_1 axis with the OXZ plane; γ is the angle between the OY_1 axis and the OX_1Y plane. Coincidence of the trihedrals $OXYZ$ and $OX_1Y_1Z_1$ is achieved by three successive rotations through the angles ψ, θ, γ around the OY, OZ and OX_1 axes respectively.

The direction cosines of the axes of the trihedral $OX_1Y_1Z_1$ relative to the orbital system of co-ordinates and the projection of the absolute angular velocity of rotation of the satellite on the axes OX_1, OY_1, OZ_1 are calculated from the equations:

$$\left.\begin{aligned}
a_{11} &= \cos\psi\cos\theta,\\
a_{12} &= \sin\psi\sin\gamma - \cos\psi\sin\theta\cos\gamma,\\
a_{13} &= \sin\psi\cos\gamma + \cos\psi\sin\theta\sin\gamma,\\
a_{21} &= \sin\theta,\\
a_{22} &= \cos\theta\cos\gamma,\\
a_{23} &= -\cos\theta\sin\gamma,\\
a_{31} &= -\sin\psi\cos\theta,\\
a_{32} &= \cos\psi\sin\gamma + \sin\psi\sin\theta\cos\gamma,\\
a_{33} &= \cos\psi\cos\gamma - \sin\psi\sin\theta\sin\gamma;
\end{aligned}\right\} \quad (3)$$

$$\left.\begin{aligned}
p &= \dot{\psi}a_{21} + \dot{\gamma} + p_1a_{11} + q_1a_{21} + r_1a_{31},\\
q &= \dot{\psi}a_{22} + \dot{\theta}\sin\gamma + p_1a_{12} + q_1a_{22} + r_1a\\
r &= \dot{\psi}a_{23} + \dot{\theta}\cos\gamma + p_1a_{13} + q_1a_{23} + r_1a
\end{aligned}\right\} \quad (4)$$

The direction cosines of the OX_1, OY_1, OZ_1 axes in the absolute system of co-ordinates are expressed from (1) and (3):

$$\left.\begin{aligned}
\bar{a}_{11} &= c_{11}a_{11} + c_{12}a_{21} + c_{13}a_{31},\\
\bar{a}_{12} &= c_{11}a_{12} + c_{12}a_{22} + c_{13}a_{32},\\
\bar{a}_{13} &= c_{11}a_{13} + c_{12}a_{23} + c_{13}a_{33},\\
\bar{a}_{21} &= c_{21}a_{11} + c_{22}a_{21} + c_{23}a_{31},\\
\bar{a}_{22} &= c_{21}a_{12} + c_{22}a_{22} + c_{23}a_{32},\\
\bar{a}_{23} &= c_{21}a_{13} + c_{22}a_{23} + c_{23}a_{33},\\
\bar{a}_{31} &= c_{31}a_{11} + c_{32}a_{21} + c_{33}a_{31},\\
\bar{a}_{32} &= c_{31}a_{12} + c_{32}a_{22} + c_{33}a_{32},\\
\bar{a}_{33} &= c_{31}a_{13} + c_{32}a_{23} + c_{33}a_{33}.
\end{aligned}\right\} \quad (5)$$

With an accuracy to terms of the first order of smallness relative to oblateness the expression for the force function determining the action of the earth's gravitational field on the satellite can be written in the form:

$$U = K\iiint\left[\frac{1}{\bar{\rho}} - \varepsilon\frac{a^2}{3\rho^3}\left(3\frac{\bar{z}_a^2}{\bar{\rho}^2} - 1\right)\right]dM, \quad (6)$$

where

$$K = Mf_a,\ \varepsilon = \alpha - \frac{m}{2},\ \alpha = \frac{a-b}{a},\ m = \frac{\Omega^2 a}{g_a},$$

$$\bar{x}_a = x_0 + x_1\bar{a}_{11} + y_1\bar{a}_{12} + z_1\bar{a}_{13},$$
$$\bar{y}_a = y_0 + x_1\bar{a}_{21} + y_1\bar{a}_{22} + z_1\bar{a}_{23},$$
$$\bar{z}_a = z_0 + x_1\bar{a}_{31} + y_1\bar{a}_{32} + z_1\bar{a}_{33},$$
$$\bar{\rho}^2 = \bar{x}_a^2 + \bar{y}_a^2 + \bar{z}_a^2$$

M_a is the earth's mass; f is the gravitational constant; M is the satellite mass; a is the equatorial radius of the earth; b is the polar radius; α is the oblateness; Ω is the angular speed of rotation of the earth; g_a is the earth's gravitational acceleration at the equator; x_1, y_1, z_1 are co-ordinates of the variable point of the satellite in the $OX_1Y_1Z_1$ system of co-ordinates; x_0, y_0, z_0 are the absolute co-ordinates of the satellite mass centre.

In calculating the integral (6) use is made of the smallness of the ratio of the characteristic linear dimension l of the satellite to the separation ρ between the mass centres of satellite and earth. Decomposing the function

$$\frac{1}{\bar{\rho}} - \varepsilon\frac{a^2}{3\rho^3}\left(3\frac{\bar{z}_a^2}{\rho^2} - 1\right)$$

into a series in the orders $\frac{x_1}{\rho}$, $\frac{y_1}{\rho}$, $\frac{z_1}{\rho}$ and disregarding terms higher than the second order of smallness in relation to unity, we obtain after integration the following equation for the force function:

$$\left.\begin{aligned}
U = {} & \frac{KM}{\rho}\left[1 - \frac{\varepsilon}{3}\cdot\frac{a^2}{\rho^2}\left(3\frac{z_0^2}{\rho^2} - 1\right)\right] - \frac{1}{4}\cdot\frac{K}{\rho^3}\left[1 - \varepsilon\frac{a^2}{\rho^2}\left(5\frac{z_0^2}{\rho^2} - 1\right)\right]\times\\
& \times\ (A+B+C) + \frac{1}{4}\frac{K}{\rho^3}\left[3 - 5\varepsilon\frac{a^2}{\rho^2}\left(7\frac{z_0^2}{\rho^2} - 1\right)\right]\cdot\left[(B+C-A)\,a_{11}^2 + \right.\\
& \left. + (C+A-B)\,a_{12}^2 + (A+B-C)\,a_{13}^2\right] + 5\varepsilon\frac{K}{\rho^3}\cdot\frac{z_0}{\rho}\cdot\frac{a^2}{\rho^2}\times\\
& \times\left[(B+C-A)\,a_{11}\bar{a}_{31} + (C+A-B)\,a_{12}\bar{a}_{32} + (A+B-C)\,a_{13}\bar{a}_{33} - \right.\\
& \left.\frac{1}{2}\varepsilon\frac{K}{\rho^3}\cdot\frac{a^2}{\rho^2}\left[(B+C-A)\,\bar{a}_{31}^2 + (C+A-B)\,\bar{a}_{32}^2 + (A+B-C)\,\bar{a}_{33}^2\right],\right.
\end{aligned}\right\} \quad (7)$$

where A, B, C are the main central moments of inertia of the satellite.

The equations of the rotational motion of the satellite in the Euler form are written

$$\left.\begin{aligned} A\frac{dp}{dt}+(C-B)\,qr&=M_{x_1},\\ B\frac{dq}{dt}+(A-C)\,rp&=M_{y_1},\\ C\frac{dr}{dt}+(B-A)\,pq&=M_{z_1}, \end{aligned}\right\} \tag{8}$$

$$\left.\begin{aligned} M_{x_1}&=(C-B)\frac{K}{\rho^3}\left\{\left[3-5\varepsilon\frac{a^2}{\rho^2}\left(7\frac{z_0^2}{\rho^2}-1\right)\right]a_{12}a_{13}+\right.\\ &\left.+10\varepsilon\frac{z_0}{\rho}\cdot\frac{a^2}{\rho^2}(a_{12}\bar{a}_{33}+a_{13}\bar{a}_{32})-2\varepsilon\frac{a^2}{\rho^2}\bar{a}_{32}\bar{a}_{33}\right\},\\ M_{y_1}&=(A-C)\frac{K}{\rho^3}\left\{\left[3-5\varepsilon\frac{a^2}{\rho^2}\left(7\frac{z_0^2}{\rho^2}-1\right)\right]a_{13}a_{11}+\right.\\ &\left.+10\varepsilon\frac{z_0}{\rho}\cdot\frac{a^2}{\rho^2}a_{13}\bar{a}_{31}+a_{11}\bar{a}_{33}-2\varepsilon\frac{a^2}{\rho^2}\bar{a}_{33}\bar{a}_{31}\right\},\\ M_{z_1}&=(B-A)\frac{K}{\rho^3}\left\{\left[3-5\varepsilon\frac{a^2}{\rho^2}\left(7\frac{z_0^2}{\rho^2}-1\right)\right]a_{11}a_{12}+\right.\\ &\left.+10\varepsilon\frac{z_0}{\rho}\cdot\frac{a^2}{\rho^2}(a_{11}\bar{a}_{32}+a_{12}\bar{a}_{31})-2\varepsilon\frac{a^2}{\rho^2}\bar{a}_{31}\bar{a}_{32}\right\} \end{aligned}\right\} \tag{9}$$

The equations of motion of the satellite mass centre in osculating elements are set out in (2) with the usual assumption regarding independent advance and rotating motion. By means of these equations the equations (2) can be simplified as follows:

$$\left.\begin{aligned} p_1&=-\varepsilon\sqrt{\frac{K}{\bar{\rho}}}\frac{a^2}{\rho^3}\sin 2i\sin u,\\ q_1&=0,\\ r_1&=\frac{\sqrt{K\bar{p}}}{\rho^2}; \end{aligned}\right\} \tag{10}$$

where $\bar{\rho}$ is the parameter of the osculating orbital ellipse.

For $\varepsilon=0$ the derived equations of rotational motion of the satellite coincide with the equations obtained in the work[9].

For the particular case of motion of the satellite mass centre along the equatorial orbit ($i=0$) the equations of rotational motion are simplified and take the form:

$$\left.\begin{aligned} A\frac{dp}{dt}+(C-B)\,qr&=(C-B)\frac{K}{\rho^3}\left[\left(3+5\varepsilon\frac{a^2}{\rho^2}\right)a_{12}a_{13}-2\varepsilon\frac{a^2}{\rho^2}a_{32}a_{33}\right],\\ B\frac{dq}{dt}+(A-C)\,rp&=(A-C)\frac{K}{\rho^3}\left[\left(3+5\varepsilon\frac{a^2}{\rho^2}\right)a_{13}a_{11}-2\varepsilon\frac{a^2}{\rho^2}a_{33}a_{31}\right],\\ C\frac{dr}{dt}+(B-A)\,pq&=(B-A)\frac{K}{\rho^3}\left[\left(3+5\varepsilon\frac{a^2}{\rho^2}\right)a_{11}a_{12}+2\varepsilon\frac{a^2}{\rho^2}a_{31}a_{32}\right], \end{aligned}\right\} \tag{11}$$

$$\left.\begin{aligned} p&=\dot{\psi}a_{21}+\dot{\gamma}+\frac{\sqrt{Kp}}{\rho^2}\,a_{31},\\ q&=\dot{\psi}a_{22}+\dot{\theta}\sin\gamma+\frac{\sqrt{Kp}}{\rho^2}\,a_{32},\\ r&=\dot{\psi}a_{23}+\dot{\theta}\cos\gamma+\frac{\sqrt{Kp}}{\rho^2}\,a_{33}. \end{aligned}\right\} \tag{12}$$

For equations (11) on a circular equatorial orbit an integral of the Yakobi type is correct:

$$\left.\begin{aligned} &\frac{1}{2}\frac{K}{\rho^3}\left(1+3\varepsilon\frac{a^2}{\rho^2}\right)\left[(C-A)\,a_{21}^2+(C-B)\,a_{32}^2\right]+\\ &+\frac{1}{2}\frac{K}{\rho^3}\left(3+5\varepsilon\frac{a^2}{\rho^2}\right)\left[(B-A)\,a_{12}^2+(C-A)\,a_{13}^2\right]+\\ &+\frac{1}{2}\left[A(\dot{\psi}a_{21}+\dot{\gamma})^2+B(\dot{\psi}a_{22}+\dot{\theta}\sin\gamma)^2+C(\dot{\psi}a_{23}+\dot{\theta}\cos\gamma)^2\right]=h. \end{aligned}\right\} \tag{13}$$

It follows from the integral (13) that relative equilibrium of the satellite on the circular orbit governed by the solution $\psi=\theta=\gamma=\dot{\psi}=\dot{\theta}=\dot{\gamma}=0$ is stable if

$$C>B,\ B>A. \tag{14}$$

The inequality (14) does not depend on ε and coincides with the stability conditions obtained in [3] and [4] considering motion in the Newtonian field of forces.

By means of the integral (13) within the stability zone (14) it is possible to obtain an evaluation for the direction cosines:

$$\left.\begin{aligned} a_{12}^2&\leqslant\frac{2h}{\frac{K}{\rho^3}\left(3+5\varepsilon\frac{a^2}{\rho^2}\right)(B-A)},\quad a_{13}^2\leqslant\frac{2h}{\frac{K}{\rho^3}\left(3+5\varepsilon\frac{a^2}{\rho^2}\right)(C-A)},\\ a_{21}^2&\leqslant\frac{2h}{\frac{K}{\rho^3}\left(1+3\varepsilon\frac{a^2}{\rho^2}\right)(C-A)},\quad a_{32}^2\leqslant\frac{2h}{\frac{K}{\rho^3}\left(1+3\varepsilon\frac{a^2}{\rho^2}\right)(C-B)}. \end{aligned}\right\} \tag{15}$$

For the earth $\varepsilon=0{\cdot}0016$ and the influence of oblateness on the evaluation does not exceed 0·5 per cent.

We will study the small oscillations of the satellite relative to the orbital system of coordinates for arbitrary location of the orbit in space. The equations of these oscillations are obtained from (8)-(9) by the usual linearisation of trigonometrical functions:

$$\left.\begin{aligned} &A\ddot{\gamma}+(C-B)\,r_1^2\gamma-(A+B-C)\,r_1\dot{\psi}-A\,\dot{r}_1\psi+f_1=\\ &=-A\dot{p}_1-2(C-B)\frac{K}{\rho^3}\frac{a^2}{\rho^2}\varepsilon c_{22}c_{33},\\ &B\ddot{\psi}+(C-A)\left(r_1^2+3\frac{K}{\rho^3}\right)\psi+(A+B-C)\,r_1\dot{\gamma}+Br_1+f_2=\\ &=(C-A)\left(p_1r_1-8\frac{K}{\rho^3}\frac{a^2}{\rho^2}\varepsilon c_{31}c_{33}\right),\\ &C\ddot{\theta}+3(B-A)\frac{K}{\rho^3}\theta+C\dot{r}_1+f_3=8(B-A)\frac{K}{\rho^3}\frac{a^2}{\rho^2}\varepsilon c_{31}c_{32}. \end{aligned}\right\} \tag{16}$$

$$
\left.\begin{aligned}
f_1 &= -(C-B)\left\{p_1 r_1 \theta + \frac{K}{\rho^3}\frac{a^2}{\rho^2}\varepsilon\left[8c_{31}c_{32}\psi - 8c_{31}c_{33}\theta - 2(c_{32}^2 - c_{33}^2)\gamma\right]\right\},\\
f_2 &= -B(\dot{p}_1 r + p_1 \dot{r}) + (C-A)\left\{-p_1^2\psi - p_1\dot{\theta} + \frac{K}{\rho^3}\frac{a^2}{\rho^2}\times\right.\\
&\qquad \left.\times\,\varepsilon\left[(5-17c_{31}^2+2c_{33}^2)\psi - 2c_{32}c_{33}\theta - 8c_{31}c_{32}\gamma\right]\right\},\\
f_3 &= C(\dot{p}_1\psi + p_1\dot{\psi}) + (B-A)\left\{p_1(\dot{\psi} - p_1\theta + r_1\gamma) - \frac{K}{\rho^3}\frac{a^2}{\rho^2}\times\right.\\
&\qquad \left.\times\,\varepsilon\left[2c_{32}c_{33}\psi - (5-17c_{31}^2+2c_{32}^2)\theta + 8c_{31}c_{33}\gamma\right]\right\}.
\end{aligned}\right\} \qquad (17)
$$

Confining the investigation of the forced solution of the system (16) within the stability zone (14) of the equilibrium position of the satellite on the circular orbit we will seek this solution in the form of serial expansion for ε:

$$
\begin{aligned}
\psi &= \psi_0 + \varepsilon\psi_1 + \ldots,\\
\theta &= \theta_0 + \varepsilon\theta_1 + \ldots,\\
\gamma &= \gamma_0 + \varepsilon\gamma_1 + \ldots.
\end{aligned}
$$

For a slightly elliptical orbit the variable expansion coefficients can in turn be written in the form of a series in degrees of eccentricity:

$$
\begin{aligned}
&\psi_0 = \psi_{00} + e\psi_{01} + e^2\psi_{02} + \ldots, \quad \psi_1 = \psi_{10} + e\psi_{11} + e^2\psi_{12} + \ldots,\\
&\theta_0 = \theta_{00} + e\theta_{01} + e^2\theta_{02} + \ldots, \quad \theta_1 = \theta_{10} + e\theta_{11} + e^2\theta_{12} + \ldots,\\
&\gamma_0 = \gamma_{00} + e\gamma_{01} + e^2\gamma_{02} + \ldots, \quad \gamma_1 = \gamma_{10} + e\gamma_{11} + e^2\gamma_{12} + \ldots.
\end{aligned}
$$

The final form the expansion with an accuracy to ε^2, εe, e^2 is written in the following manner:

$$
\left.\begin{aligned}
\psi &= \psi_{00} + e\psi_{01} + \varepsilon\psi_{10},\\
\theta &= \theta_{00} + e\theta_{01} + \varepsilon\theta_{10},\\
\gamma &= \gamma_{00} + e\gamma_{01} + \varepsilon\gamma_{10}.
\end{aligned}\right\} \qquad (18)
$$

The first two coefficients of the expansion (18) are calculated in the work[4]

$$
\psi_{00} = \theta_{00} = \gamma_{00} = 0,\ \psi_{01} = \gamma_{01} = 0,\ \theta_{01} = \frac{2C}{3(B-A)-C}\sin\omega_0 t.
$$

Here ω_0 is the mean angular velocity of motion of the satellite mass centre along the elliptical orbit. The functions ψ_{00}, θ_{00}, γ_{00} correspond to the equilibrium position of relative satellite equilibrium in the Newtonian field of forces ($\varepsilon=0$) on the circular orbit ($e=0$). The functions ψ_{01}, θ_{01}, γ_{01} determine the eccentricity oscillations of the satellite in the plane of the elliptical orbit.

The functions γ_{10}, θ_{10}, ψ_{10} represent the forced solution of the following system of equations

$$\left.\begin{aligned}
&A\ddot{\gamma}_{10}+(C-B)\,\omega_0^2\gamma_{10}-(A+B-C)\,\omega_0\dot{\psi}_{10}=\\
&=(A+B-C)\,\omega_0^2\,\frac{a^2}{\rho^2}\sin 2i\cos\omega_0 t,\\
&B\ddot{\psi}_{10}+4\,(C-A)\,\omega_0^2\psi_{10}+(A+B-C)\,\omega_0\dot{\gamma}_{10}=\\
&=-5(C-A)\,\omega_0^2\,\frac{a^2}{\rho^2}\sin 2i\sin\omega_0 t,\\
&C\ddot{\theta}_{10}+3\,(B-A)\omega_0^2\theta_{10}=\\
&=4\,(B-A)\,\omega_0^2\,\frac{a^2}{\rho^2}\sin^2 i\sin\omega_0 t.
\end{aligned}\right\}\qquad(19)$$

In calculating ψ_{10}, θ_{10}, γ_{10} on the right-hand side of equations (19) it should be considered that $e=\varepsilon=0$.

Having determined ψ_{10}, θ_{10}, γ_{10} from (19) we obtain the equations for ψ, θ, γ:

$$\left.\begin{aligned}
&\psi=-\varepsilon\,\frac{4\,(C-A)+B}{3\,(C-A)}\,\frac{a^2}{\rho^2}\sin 2i\sin\omega_0 t,\\
&\theta=\varepsilon\,\frac{4\,(B-A)}{3\,(B-A)-4C}\,\frac{a^2}{\rho^2}\sin^2 i\sin 2\omega_0 t+e,\ \frac{2C}{3\,(B-A)-C}\sin\omega_0 t,\\
&\gamma=\varepsilon\,\frac{(C-A)+B}{3\,(C-A)}\,\frac{a^2}{\rho^2}\sin 2i\cos\omega_0 t.
\end{aligned}\right\}\qquad(20)$$

With sufficient removal from the boundaries of the stability zone of the trivial solution on the circular orbit (C>B>A) the influence of ε on the forced solution is slight.

The table sets out the amplitudes of the angles ψ, θ, γ due to oblateness as a function of $\theta_A=\dfrac{A}{C}$ for $i=65°$ and $\theta_B=\dfrac{B}{C}=0{\cdot}9$. In the calculations $\dfrac{a}{\rho}$ was taken as equal to unity. It is seen from the table that over a wide range of variation of θ_A the amplitude of angle fluctuations does not exceed 12′.

θ_A	...	0·8	0·7	0·6	0·5	0·4	0·3	0·2	0·1
ψ, min	...	−11·9	−9·8	−8·8	−8·2	−7·7	−7·4	−7·2	−7·0
θ, min	...	−0·5	−1·1	−1·8	−2·6	−3·6	−4·9	−6·7	−9·0
γ, min	...	7·7	5·6	4·6	3·9	3·5	3·2	3·0	2·8

Resonance phenomena are not possible in the first terms of equations (20) within the zone $C>B>A$ since $C-A>0$ and $3(B-A)-4C<-C$.

The second term in the expansion for the angle θ, governed by the elliptical orbit, is investigated in the work[4].

With $A=B=C$ the satellite rotates with constant angular velocity relative to a stationary axis in absolute space. This case cannot be described by the linearised equations (16).

The rotational motion of the satellite relative to the absolute system of co-ordinates is determined by the relations (5).

A more precise forced solution (20) by introducing higher powers of e is not difficult. Further terms of the expansion for ε do not enter into consideration for the purposes of the present problem.

REFERENCES

1. F. R. Mul'ton, *Introduction to Celestial Mechanics*. Onti, Moscow, 1935. (In Russian).
2. D. E. Okhotsimskii, T. M. Eneyev and G. P. Taratynova, *U.F.N.*, **63**, No. 1a. 33 (1957).
3. F. Tisserand, *Traité de Mécanique Celéste*, II, Paris (1891).
4. V. V. Beletskii, *Isk. Sput. Zem.*, **3**, 13 (1959).

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 179–182. Printed in Breat Britain.

VARIATION IN THE POSITIVE ION CONCENTRATION WITH ALTITUDE FROM DATA OF MASS SPECTROMETRY MEASUREMENTS ON THE THIRD SATELLITE

V. G. ISTOMIN

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 6, p. 127, (1961)

The amplitude (intensity) of ion current peaks of different masses, recorded by the mass spectrometer on a satellite, is a function of several variables and naturally must depend on the concentration of the respective ions. Assuming that the thermal velocity of ions is much smaller than the speed of the satellite this relationship can be written as follows:

$$i_M = q \cdot n_M \cdot v_s \cdot S_{eff}(\theta, M, \phi) \cdot \eta(M, \phi, v_s), \qquad (1)$$

where i_M is the amplitude of the ion current peak for the mass number M on the mass spectrogram, q is the ion charge, n_M is the ion concentration with mass number M, v_s is the speed of the satellite, S_{eff} is the "effective cross-section" of the entry orifice of the mass spectrometer tube and η is the "current efficiency" of the instrument, that is to say the ratio of the ion current of mass M entering the analyzer to the current of the same ions on the collector i_M. Hence to determine the ion concentration it is necessary to know the "current efficiency" of the mass spectrometer and the effective cross-section S_{eff}. These two values are a function of the angle θ between the tube axis and the velocity vector (i.e. the tube orientation relative to the velocity vector of the satellite), of the absolute velocity v_s of the potential ϕ to which the satellite is charged relative to the unperturbed plasma and finally of the mass number of the ions M. This latter value is measured directly with the mass spectrometer. If the mass spectrometer were to afford also a measurement of the satellite potential ϕ and the angle θ then, knowing the effective cross-section of the entry orifice of the tube as a function of θ and ϕ for different mass numbers M, it would be possible to obtain from the mass spectrometer measurements not only the composition but also the ion concentrations of different masses in absolute units.

In setting up the experiment it was not intended to obtain absolute or relative positive ion concentrations and their relationship to altitude but the treatment of the data obtained indicated that ion concentration in relative units as a function of altitude could be clearly demonstrated from the mass spectrometer data. It follows actually from equation (1) that for ions of a given mass M the amplitude of the corresponding ion current peak on the mass spectrogram will be proportional to their concentration n_M and will be also a function of the satellite potential ϕ and the angle θ. The variations in the peak amplitudes with mass number M, recorded with the same satellite potential and the same orientation of the mass spectrometer tube will be proportional to the variation in the ion concentration of the given mass.

As shown by spectral analysis the satellite potential was negative, was approximately $3V$[1,2] and did not vary appreciably. The potential was estimated from the shift of the mass peaks in the spectra[2], whilst its variations can be assessed from the variations of the relative intensities of the spurious (harmonic) peaks occurring in the spectra. It is known from the theory of the radio frequency mass spectrometer that the relative intensity of harmonic peaks is a strict function of the retarding potential level and may consequently be a sensitive indicator of the constancy of the effective retarding potential. Since the satellite potential is a component of the effective retarding potential then the intensity of the harmonic peaks must vary sharply with variation in the satellite potential. Measurements of relative peak intensity of the "light harmonic" of atomic

oxygen have shown that for more than 80 per cent of all the spectra it has a value $0{\cdot}1 \pm 0{\cdot}2$. On the basis of this it can be regarded that the satellite potential on the sections of the orbit of interest here did not vary appreciably and all the mass spectrometer measurements were carried out with some fixed satellite potential φ_0, for which variations can be disregarded to a first approximation*.

The orientation of the mass spectrometer tube relative to the velocity vector was determined from magnetometer measurement data[4]. The intensities of the atomic oxygen peak, obtained with different angles θ were reduced to $\theta = 0°$. These reduced intensities of the 0^+ peak are shown in Fig. 1 as a function of the altitude of flight of the satellite. For each point on the curve, discrepancies are shown associated with the possible error in determining the angle θ†. The points indicated by different symbols correspond to measurements on different loops. The table gives the dates, loop number and local time for the point of satellite projection on the earth's surface at the moment of its location in the zone of the F_2 layer maximum.

Date	Loop No.	Local time	
		h	*m*
May 18	42	14	00
May 19	56	15	30
May 21	81	10	30
May 23	109	11	00

The data for 19th and 23rd May are plotted on the curve without any further normalization whilst the data for the 8th and 21st May are normalized for the altitude zone around 300 km; the normalizing factors are respectively 2·6 and 1·66.

It is seen from the curve that the atomic oxygen ion peak intensity reduced to $\theta = 0°$, proportional to the 0^+ ion concentration, points to a reduction in concentration with increasing altitude. In the altitude range 300-550 km the data for the four days coincide fairly well, pointing to a reduction in concentration to approximately half at an altitude 550 km as compared with the zone of the F_2 layer maximum (approximately 300 km). For higher altitudes the data diverge considerably for the different days, which apparently points to the relative instability from day to day of the 0^+ ion concentration trend with altitude.

Considering a mean relationship for the measurement data it can be seen that up to an altitude of 900 km the 0^+ ion concentration falls off less than one order of magnitude as compared with the zone of the F_2 layer maximum.

Mass spectrum measurements of ions in the range from 6 to 48 atomic mass units have shown that in addition to the atomic oxygen ion which predominates at altitudes above 225 km there are in the ionosphere also molecular ions N^+ [1,2]. Above 300 km however the total concentration of all these ions is less than 10 per cent of the atomic oxygen ion concentration. It can be considered therefore that the points plotted on the diagram reflect (with an error not exceeding 10 per cent) the altitude trend of the concentration of all the positive ions with mass numbers from 14 to 32 atomic mass units.

The altitude trend of positive ion concentration obtained can be compared with the altitude trend of electron concentration determined from radio observation data on the first and third Soviet artificial satellites (curves 1 and 2 in Fig. 1). Curve 1 was obtained from observations of "radio ascent" and "radio descent" of the first artificial earth satellite whilst, as shown by the authors[5], the electron concentration values for altitudes higher than 600 km are possibly exaggerated. The curve 2 was obtained from observations of the Faraday effect for

* This conclusion cannot be regarded as unexpected (see for example [3]). It is recalled that the mass spectrometer measurements were carried out only in the zone of the ionosphere illuminated by the sun and were not carried out in the zone of the earth's shadow.

† An assessment of the error in determining the angle yields the value ±10 per cent.

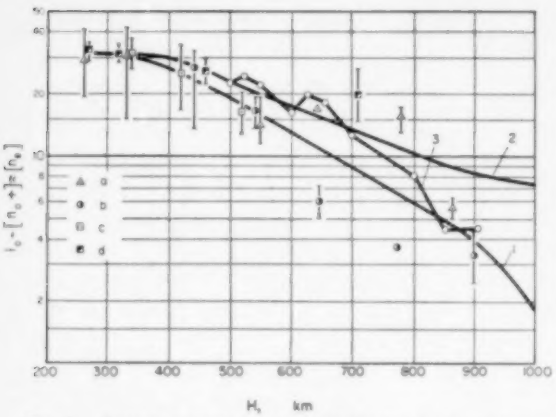


Fig. 1. Atomic oxygen ion peak intensity i_0+ positive ion concentration (n_0+) and electron concentration (n_e) as a function of altitude H.
i_0+ measurements: a - 18th May; b - 19th May; c - 21st May; d - 23rd May (see Table 1);
1 - electron concentration from data of "radio ascent observations" and "radio descent" observations of the first satellite [5]; 2 - from data of observations of the Faraday effect (along the loops of the third artificfal earth satellite) [6]; 3 - positive ion concentration from measurements on the third artificial earth satellite 19th May (loop 56) [7] (all values in relative units).

signals of the "Beacon" transmitter installed on the third artificial earth satellite from data of 135 passages of the satellite in the period from 16th May to 30th September 1958[6]. The data are comparable in the sense that the measurements were carried out along satellite orbits approximately in the same season (summer-autumn) and in the middle latitudes of the northern hemisphere. In plotting curves 1 and 2 on the graph they were superposed as in the case of the mass spectrometer measurement data in the altitude zone around 300 km although the absolute electron concentration values in the zone of the F_2 layer maximum indicated by the authors [5] and [6] diverge more than two times and are equal respectively to $1{\cdot}8 \times 10^6$ and $0{\cdot}7 \times 10^6$ cm^{-3}.

A comparison of the data shows that up to an altitude of 550 km the variation in the positive ion concentration shows satisfactory agreement with the variation in the electron concentration whilst above 650 km the data of the mass spectrometer measurements for three different days reveal a considerable dispersion; the 0^+ ion peak intensities for different days differ sharply from each other but on average at altitudes 650-900 km the points are located around the curves of electron concentration variation. This is quite regular since electron concentration data are averaged from several measurements and cannot reflect day to day variations.

The altitude - positive - ion - concentration relationship obtained from the mass spectrometer data can also be compared with experimental results for the determination of positive ion concentration employing ion traps, which as is known were installed on the third satellite. The diagram shows mean concentration measurement data on 19th May (curve 3) taken from the graph produced in the work[7]. It is seen that the curve 3 agrees quite well both with the mean altitude -0^+ ion concentration relationship and with the altitude trend of electron concentration.

The agreement between the measured N^+, 0^+, N_2^+ NO^+ and 0_2^+ positive ion concentration altitude trend and the electron concentration trend must serve as a pointer to the fact that light ions with mass numbers outside the range of the instrument (H^+, H_2^+ hydrogen ions and He^+ helium ions), do not exist in appreciable quantities and the ionosphere remains oxygen-nitrogen up to an altitude of 900 km. Actually with the existence of light ions it is natural to expect an increase in their relative concentrations with altitude, similar to what is observed for atomic nitrogen ions in relation to atomic oxygen ions and for atomic oxygen ions in relation to still heavier ions[1,2]. An increase in the relative concentration of H^+, H_2^+ and He^+ ions would lead to a divergence between the electron and ion concentration values.

Thanks are accorded to B. A. Mirtov and K. I. Gringauz for interest in the work and consideration of the results, together with V. V. Beletskii and Yu. V. Zonov for the submission of data on satellite orientation.

REFERENCES

1. V. G. Istomin, *Dokl. AN.S.S.S.R.*, **129**, 81 (1959).
2. V. G. Istomin, *Isk. Sput. Zemli*, **4**, 171 (1960). In Press.

3. V. G. KUST and V. I. MOROZ, *Isk. Sput. Zemli*, 7 (1961). In Press.
4. V. V. BELETSKII and YU. V. ZONOV, *Isk. Sput. Zemli*, 7 (1961). In Press.
5. YA. L. AL'PERT, F. F. DOBRYAKOVA, E. F. CHUDENSENKO and B. S. SHAPIRO, *U.F.N.*, **65**, 161 (1958).
6. L. KLINKER, K. K. KNUT and K. H. SCHMELOVSKII, *Z. Met.*, **13**, 192 (1959).
7. K. I. GRINGAUZ, V. V. BEZRUKIKH and V. D. OZEROV, *Isk. Sput. Zemli*, **6**, 63 (1961).

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 183–193. Printed in Great Britain.

MEASUREMENTS OF ELECTRON CONCENTRATION IN THE IONOSPHERE UP TO ALTITUDES OF 420-470 km, CARRIED OUT DURING THE INTERNATIONAL GEOPHYSICAL YEAR BY MEANS OF RADIO-WAVE TRANSMISSIONS FROM THE GEOPHYSICAL ROCKETS OF THE ACADEMY OF SCIENCES U.S.S.R.

K. I. GRINGAUZ and V. A. RUDAKOV

Translated by R. MATTHEWS from *Iskusstvennye Sputniki Zemli*, No. 6, p. 48 (1961)

1. INTRODUCTION

Since 1954 a study has been carried out in the Soviet Union, in conjunction with the vertical launchings of geophysical rockets, of the altitude distribution of free electron concentration in the ionosphere—one of the most important physical characteristics of the ionosphere from the standpoint of radio-wave propagation. The results of these measurements have been published in part[1-2] and reproduced in [3]. The instrumentation employed for the measurements has been described in a paper published in this journal [4].

Despite the recent success in the study of the ionosphere employing instruments installed on artificial earth satellites (see for example [5]), investigations employing vertically launched rockets do not and cannot lose their importance. This is due to two things. The first is the fact that artificial satellites are not as a rule launched to altitudes below approximately 200 km (since otherwise they have a short life) and hence measurements are possible in the range below 200 km only employing rockets. The second and most important factor is that the horizontal velocity of a satellite is usually much greater than its vertical velocity and therefore variations in the investigated characteristic of the medium associated with changes in the measuring equipment with altitude are difficult to separate from the variations associated with the geophysical (latitude and longitude) changes.

The present paper examines a number of problems relating to methods for measuring radio-wave dispersion and the Faraday effect, as employed in the determination of electron concentration n_e, setting out results of measurements carried out during the launching of the three geophysical rockets to altitudes 450-470 km in 1958. It was possible from these measurements, which were carried out at different times of the day and year, to determine the n_e distribution with altitude in the ionosphere region including part of what is termed the outer ionosphere (located above the F-layer maximum and completely inaccessible to study by the normal radiosonde methods of ionosphere investigation from the ground). The measurements were carried out over one and the same geographical point and by the same method so that the comparison of results presents obvious interest.

2. THE DISPERSION INTERFEROMETER METHOD

The relationship of the speed of propagation of radio-waves to the frequency (the radio-wave velocity dispersion) in the ionosphere is governed substantially by the free electron concentration n_e. Hence radio-wave velocity dispersion measurements can be used for determining n_e.

One of the forms of radio interference instruments proposed by L. I. Mandel'shtam and N. D. Papaleski is particularly suitable for measuring radio-wave velocity dispersion, being a version of the phase probe method called by them dispersion interferometer[6]. The phase probe method was first used for the

study of the ionosphere at the time of the 1936 solar eclipse[7].

In the case of the dispersion interferometer, coherent radio-waves are emitted from the point A at frequencies f_1 and f_2 ($f_1 = pf_2$, where $p = \frac{m}{n} > 1$, m and n are whole numbers; in the particular case $n = 1$). The radio-waves are received at point B and their phase difference $\Delta\phi$ is recorded, being the phase difference referred to the higher frequency i.e. $\Delta\phi = \phi_1 - p\phi_2$. At the point B

$$\Delta\phi = \frac{2\pi p f_2}{c}\left(\int_0^L n_1(l)\,dl - \int_0^L n_2(l)\,dl\right), \qquad (1)$$

where L is the distance between A and B; integration is carried out along the path from the transmitting to the receiving antenna; n_1 (1) and n_2 (1) are coefficients of refraction of the radio-waves for frequencies f_1 and f_2 respectively.

For sufficiently short radio-waves the coefficient of refraction of the ionosphere can be taken as

$$n(f) \approx 1 - \frac{e^2}{2\pi m f^2} n_e, \qquad (2)$$

where n_e is the electron concentration, e and m are the electron charge and mass and f is the oscillation frequency. All the values in (2) and subsequently are in the gauss system of units.

The expression (2) derives from the known expressions for the coefficient of refraction of radio-waves (see for example [8]) assuming that n differs slightly from unity and that electron collision and the influence of the earth's magnetic field can be disregarded. For this purpose the condition must be satisfied

$$\nu \ll \omega = 2\pi f$$

and

$$\omega_H \ll \omega,$$

where ν is the electron collision frequency (for altitudes higher than 70 km $\nu \leqslant 10^7$ sec^{-1} [9]), $\omega_H = 8{\cdot}82 \times 10^6$ sec^{-1} is the gyromagnetic frequency. Using the expression (2) for frequencies $f \geqslant 20$ Mc/s yields a relative discrepancy $\frac{\delta n}{n} \leqslant 1$ per cent.

Employing (2) expression (1) can be written in the form

$$\Delta\phi = \frac{e^2}{cmf_2}\left(\frac{p^2-1}{p}\right)\int_0^L n_e(l)\,dl = K\int_0^L n_e(l)\,dl. \qquad (3)$$

Considering that the values n_e and L vary in time (that is to say the point A moves relative to the point B) then during the time Δt during which the point A moves through ΔL, the variation in the phase difference of the oscillations received at point B will be

$$\Delta\Phi = K\left[\int_L^{L+\Delta L} n_e(l)\,dl + \left(\int_0^L \frac{\partial}{\partial t} n_e(l)\,dl\right)\Delta t\right] = \Delta\Phi_{\text{loc}} + \Delta\Phi_{\text{int}} \qquad (4)$$

(it is assumed that during the time Δt the value $\int_0^L \frac{\partial}{\partial t} n_e(l)\,dl$—the rate of variation of the integral concentration—varies very slightly).

Hence the recorded increase in the phase difference can be regarded as consisting of two components:

(a) $\Delta\Phi_{\text{loc}}$, associated with increase in path L through time Δt and depending on the electron concentration on the section ΔL and

(b) $\Delta\Phi_{\text{int}}$, associated with variation over the time Δt of the integral content of electrons in the column from the observer to the commencement of the section ΔL.

For a given rate of variation, in time, of the integral concentration $\int_0^L n_e(l)\,dl$ the specific rate $\Delta\Phi_{\text{loc}}$ in (4) is the greater the greater is n_e on the section ΔL and the greater is the speed of movement of the point A (or the smaller Δt, corresponding to the given value ΔL and consequently the smaller $\Delta\Phi_{\text{int}}$).

In the case of a stationary point A (this was actually the case at the time of the 1936 solar eclipse when the points on the ground A and B were stationary and variations in the optical paths of radio-waves reflected from the ionosphere were studied) the first component is

absent and the value $\Delta\Phi$ is determined by the mean speed of variation $\int_0^L n_e(l)\,dl$ during the time Δt. If the point A is a rocket launched in a direction approaching the vertical, then the same situation occurs at the peak of the trajectory where the vertical velocity of the rocket changes its sign, passing through zero.

It is evident that in measuring $\Delta\Phi$ at the time when the rocket is near the peak of the trajectory ($h \approx h_{max}$) the local value n_e cannot be determined but it is possible to estimate $\int_0^L \frac{\partial}{\partial t} n_e(l)\,dl$.

We will indicate the integral concentration of electrons in the vertical column of air with altitude h by $N_e = \int_0^h n_e dh$ and substitute, for the case of a vertically launched rocket and a receiving point B near to the projection of the peak of the trajectory on the ground, the distance O between the observer and rocket with rocket altitude h above the ground.

Quantitative evaluations are carried out below of the rate of variation of N_e, but at the moment we will remark only that already at some kilometres from the peak of the rocket trajectory $\Delta\Phi_{loc}$ substantially exceeds $\Delta\Phi_{int}$. At the same time for determining the mean value of n_e over the altitude range Δh use can be made of the following expression obtained from (4) and (3),

$$n_e = \frac{L}{K\Delta h}\int_h^{h+\Delta h} n_e(h)\,dh = \frac{cmf}{e^2}\,\frac{p}{p^2-1}\,\frac{\Delta\Phi}{\Delta h} \qquad (5)$$

(disregarding $\Delta\Phi_{int}$ as was done in [1] and [2]).

The interval of time during which the rocket moves in altitude through Δh and during which the corresponding change $\Delta\Phi$ is recorded will be called the "averaging time" shown by Δt_i.

Substituting in (5) the shifts along the connecting direction L by the shifts in altitude h introduces an error in determination of n_e, diminishing with increasing h on the basis of geometrical concepts. For altitude $h = 100$ km and for a removal of the receiving station from the projection of the peak of the rocket trajectory on the ground equal to 5 km the relative discrepancy in determining n_e due to the inequality $\Delta h \neq \Delta L$ is $\frac{\partial n_e}{n_e} \approx 0{\cdot}5$ per cent.

It is appropriate to observe the following concepts forming the basis of the choice of radio-wave lengths for these experiments carried out during the I.G.Y.

Since it was envisaged that the investigations would proceed partly in the outer ionosphere, the radio frequencies must certainly exceed any maximum critical frequencies of the ionosphere layers encountered, in order to maintain uninterrupted communication of the rocket with the ground at all altitudes.

When employing radio-waves in the ultra-short wave range, reflections of the radio-waves from the ionosphere zone above the rocket are impossible (in any case for vertical propagation of the radio-waves), and therefore reflected interference signals could not arrive at the receiving stations. Using ultra-short waves makes it possible to employ highly directional antennae which would also considerably reduce the possibility of receiving interference signals (in this case reflected from the earth's surface). In addition the use of radio-waves in the ultra-short wave range affords ionosphere characteristics averaged over smaller zones than when using longer radio-waves since the size of the zone participating in forming the signal is proportional, to a first approximation, to the area of the first Fresnel zone (i.e. the wavelength).

These concepts led to a choice of radio-waves in the ultra-short wave range with lengths $\lambda \approx 2$ and 6 m (frequency $f_1 = 144$ Mc/s and $f_2 = 48$ Mc/s) as the principal waves and $\lambda = 12$ m (frequency $f_3 = 24$ Mc/s) as the supplementary wave.

During the I.G.Y. experiments in ionosphere research were carried out, in addition to the Soviet Union, in the U.S.A. employing radio-waves emitted from rockets under the direction of Seddon and Jackson[10]. The experiments were carried out in the Arctic zone (Fort Churchill, Canada, 59° northern latitude) yielding very interesting results relating to the ionos-

phere zone below the F layer maximum. In these experiments, as in the preceding experiments by the same authors[11, 12], a method was employed differing from that described above in that, instead of the phase difference of coherent radio-waves, the difference in the Doppler shifts of their frequencies was recorded (being physically equivalent). It is interesting to note that the Seddon and Jackson measurements were carried out employing much longer wavelengths ($f_z = 7{\cdot}75$ Mc/s, $p = 6$). If such long waves had been employed on our equipment it would have been impossible to carry out measurements at altitudes above 300 km since the critical frequencies of the F layer amounted to 14 Mc/s and radio reception of the signals from the rocket would have been impossible.

3. METHOD OF DETERMINING THE ELECTRON CONCENTRATION IN THE IONOSPHERE FROM MEASUREMENTS OF THE FARADAY EFFECT

In addition to measurements of radio-wave dispersion, requiring the recording of the phase difference of two radio-waves with different frequencies, for the purpose of determining the electron concentration at different altitudes use can be made of observations, at the reception point, of the rotation of the polarization plane of radio-waves emitted from the rocket travelling vertically in the ionosphere. These determinations were carried out during the time of the experiments described in the present article; at the same time, measurements of rotation of the polarization plane of the received radio-waves (Faraday effect) were undertaken simultaneously with the measurements of the velocity dispersion of the same radio-waves using the same receiving equipment[4]. This rendered the results of both measurements readily comparable.

A very material circumstance providing ready and simple interpretation of the measurements of the rotation of the polarization plane was that the geophysical rockets of the Academy of Sciences U.S.S.R. launched to altitudes over 400 km were practically fully stabilized throughout their free flight. This eliminates the possibility of rotation of the polarization plane of the received radio-waves as the result of the rotation of the transmitting antennae installed on the rocket with linear polarization.

It is known that radio-waves propagating in the ionosphere are split under the action of the earth's magnetic field into two components having different phase velocities. As the result of this, at the reception point the plane of polarization of the radio-waves emitted from the ionosphere with linear polarization is turned through the angle θ. Using the known formulae of the theory of radio-wave propagation in the ionosphere, taking into account the presence of the geomagnetic field (for example, [3]), it can readily be shown that in the case of arbitrarily directed radio-wave propagation relative to the magnetic field strength vector, the polarization plane turns (while travelling along a line of distance ΔL) through the angle

$$\theta = \frac{e^3}{2\pi c^2 m^2} \frac{1}{f^2} \int_L^{L+\Delta L} H_l n_e(l)\, dl, \qquad (6)$$

where H_L is the magnetic field component along the direction of propagation, whilst f is a frequency sufficiently high that absorption in the ionosphere is small and the coefficients of refraction of both components are near to unity.

During reception of radio-waves emitted from the rocket by the radio transmitters of the dispersion interferometer, records were made of the voltages at the inputs of the receivers having antennae with linear polarization. On these records the rotation of the polarization plane of the received signals is recorded in the form of periodical fading of the signals. The minimum signal levels (corresponding to the moments when the polarization plane of the received radio-waves and the polarization plane of the receiving antenna are mutually perpendicular) are particularly clearly established. The distances between the two minima correspond to variations in the angle θ by π. Fig. 1 shows part of the graph of the voltage variations at the inputs of one of the receivers connected to receiving antennae with identical polarization. The graph was constructed from data obtained at the time of launching the

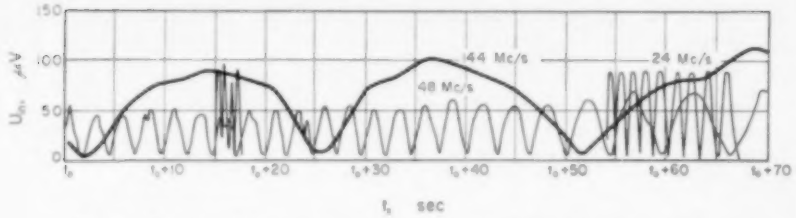


Fig. 1. Section of the curves of variation, in time, of the signal levels received during the rocket flight 27th Aug 1958 (the curve for $f_2 = 24$ Mc/s is not complete).

Academy of Sciences U.S.S.R. rocket on 27th August 1958; it clearly brings out the relation $\theta \sim \frac{1}{f^2}$, leaving no doubt that the recorded fluctuations in the signal levels were due to the Faraday effect.

It follows from (6) that if $\theta = \pi$ then taking the value H_L as known and equal to the vertical component of the geomagnetic field H_B (not varying along the path ΔL) and substituting ΔL by Δh (see section 2) the mean electron concentration of the given section of the path can be determined as

$$n_e = \frac{2\pi^2 c^2 m^2}{e^3} \frac{f^2}{\Delta h} \frac{1}{H_B} \qquad (7)$$

The value Δh is determined from the times t_1 and t_2 corresponding on the receiver input voltage record to two adjacent polarization minima since the rocket co-ordinates are known at each moment in time. Fig. 2 shows the length of the sections of the rocket path in the ionosphere as a function of mean electron concentration during the passage of which rotation of the polarization plane of the received radio-waves with frequencies 24, 48 and 144 Mc/s is equal to π (it is assumed that $H_B = 0{\cdot}4$ eV along these paths).

Employing at the receiving station two receivers with independent antennae, having mutually perpendicular polarization, allows the reduction by half of the altitude intervals for which n_e (1) is averaged. The altitude intervals over which the value n_e is averaged are substantially greater than when using the phase dispersion method set out in the previous section. However the method now being considered affords a clear representation of the altitude trend of electron concentration in the ionosphere if a fully stabilized rocket is employed. This method is very attractive due to its outstanding simplicity and the possibility of utilizing radio-waves of one frequency.

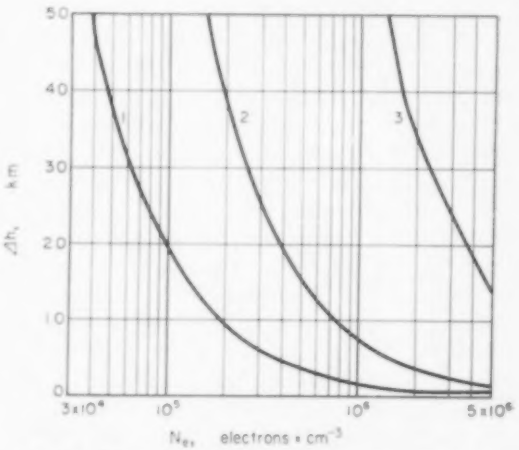


Fig. 2. The altitude interval Δh through which the polarization plane of the radio-waves rotates through angle π as a function of n_e.
1—for frequency 24 Mc/s; 2—for frequency 48 Mc/s; 3—for frequency 144 Mc/s.

4. MEASUREMENT RESULTS

As shown in the description of the instrumentation employed[4], recording of the measurement results of radio-wave velocity dispersion was carried out by two methods—recording the interference curve at the output of the phase detector of the receiver and recording the displacements of the characteristic points of the Lissajous figures on the oscillograph tube. Both methods of recording

yield practically coinciding results; for the sake of determinacy only the first method will be intended in the rest of this paper.

The oscillation frequency on the recorded interference curve depends on the electron concentration at the corresponding section of the rocket trajectory and on the speed of motion of the rocket on this section. Hence the form of the record during the time of the rocket flight varies appreciably from section to section. Fig. 7 in the paper[4] (p. 42) shows some sections of film corresponding to different altitudes of flight. This film shows the initial experimental results the development of which yields the determination of n_e at different altitudes.

Results are set out below of electron concentration measurements carried out with Academy of Science U.S.S.R. geophysical rocket launchings to altitudes above 400 km in 1958. All three experiments were carried out at the same geographical location in the middle belt of European U.S.S.R.

Figure 3 shows curves of n_e against altitude h obtained by the dispersion interferometer method (using frequencies 144 and 48 Mc/s) on sections of the trajectories corresponding to rocket ascents. Calculations of n_e were carried out from equation (5) taking account of the frequency (phase) multiplication coefficient of the receiver. For the reasons stated in the following section, the averaging time in calculating the n_e values from the disperson interferometer data was taken as constant for all altitudes ($\Delta t_i \approx 0{\cdot}2$ sec), as the result of which the altitude intervals over which the mean values n_e (h) were calculated diminished as the peaks of the rocket trajectories were approached. The curves shown in Fig. 3 are smooth within ± 5 per cent of the n_e value.

As mentioned previously the experiments described were carried out with fully stabilized rockets. Hence from the receiver input voltage records it was possible to construct n_e distributions employing the relationship of the angle of rotation of the polarization plane (Faraday effect) to electron concentration. Fig. 4 shows one of these distributions constructed from data

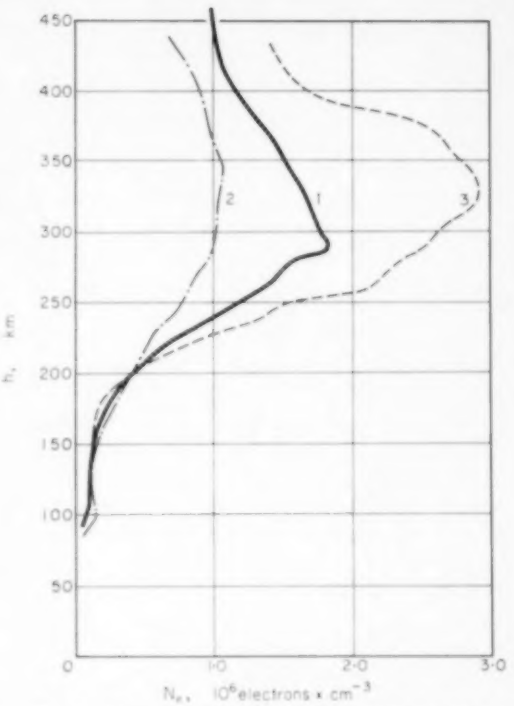


Fig. 3. Curves of electron density against altitude, obtained by the dispersion method.
1—21st Feb. 1958, 11h 40m; 2—27th Aug. 1958, 8h 06m; 3—31st Oct. 1958, 15h 54m.

of the experiment of 27th August 1958. The length of each vertical section corresponds to rotation of the polarization plane of radio-waves with frequency 48 Mc/s through π and is the altitude interval over which averaging occurs in determining n_e.

In the region of the top point of the trajectories of rockets launched 21st February and 27th August 1958 the rates of variation of integral concentration were determined $\frac{\partial}{\partial t} N_e = \int_0^{h_{max}} \frac{\partial}{\partial t} n_e(h)\, dh$ (see section 2). This value was determined as $\frac{1}{K}\left(\frac{\Delta\Phi}{\Delta t_i}\right)$ near to the t value corresponding to $v_h \approx 0$, with $\Delta t_i \approx 0{\cdot}3$ sec. In both cases the values $\frac{\partial}{\partial t} N_e$ were near to 5×10^9 electron/cm^3/s. At the

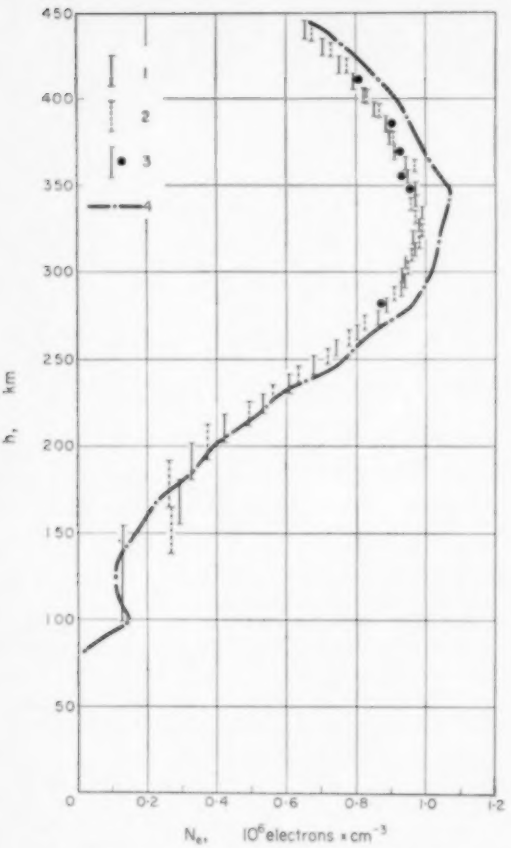


Fig. 4. n_e as a function of h, obtained from measurements of the Faraday effect 27th Aug. 1958. 1—from coincident data of 2 points; 2—from data of the third point; 3—coincident values on 3 points; 4—curve obtained by the dispersion method at the same time (on radio-waves with frequencies 144 and 148 Mc/s).

time of the third test (31st October 1958) the performance of the dispersion interferometer instrumentation was interrupted in the region of the trajectory peak, as a result of which the investigation of the variation in the integral concentration was not accomplished.

At the time of the rocket launching, radio soundings of the ionosphere were carried out employing a panorama ionosphere station near the launching site. A series of altitude-frequency curves taken with a frequency of four times per minute simultaneously with the rocket flight provides a comparison of the rocket measurement results with ionosphere station data together with the possibility of assessing the stability during the measurements of the state of the ionosphere zone below the F-layer maximum.

A study of these characteristics shows that during the time of the rocket flights (approximately 10 minutes) appreciable variations in the critical frequencies and corresponding altitudes did not occur. The minimum variation in critical frequency which could be observed on the altitude-frequency curves is approximately 0·2 Mc/s. Assuming that this variation occurred during 10 minutes then it is possible to determine the upper limit of possible critical frequency variation rate and (knowing the critical frequency) the rate of variation in the concentration N_e in the ionisation maximum. According to ionosphere station data during the experiments the relative rate of variation of the critical frequency of the F layer did not exceed 3×10^{-5} s^{-1}, corresponding to a rate of variation of the integral concentration $N_e \frac{\partial}{\partial t} N_e \approx 1{\cdot}5 \times 10^9$ electrons cm^2 s; it was assumed in the calculation that the critical frequency of the F layer f_{cr}, $F = 12$ Mc/s, $N_e \int_0^{h_F} n_e(h)\, dh \approx 2 \times 10^{13}$ electrons cm^2 (h_F is the altitude of the F-layer maximum), and the relative variations of n_e are the same at all altitudes.

5. ERRORS IN THE DETERMINATION OF ELECTRON CONCENTRATION

We will evaluate the error associated with the fact that in constructing the curves for n_e against altitude h the calculations are carried out for known values of $\Delta\Phi$ and Δh disregarding $\Delta\Phi_{int}$ (see [4] and [5]). We will start from the basis that the rate of variation of integral concentration $\frac{\partial}{\partial t} N_e$ over the whole ionosphere section of the rocket trajectory passing through the outer ionosphere is equal to a definite value near to the peak 5×10^9 electrons/cm^2/s. This

value exceeds the possible value $\frac{\partial}{\partial t} N_e$ obtained from the ionosphere station data for the lower section of the F layer and, as shown in section 6, appreciably exceeds all the previously published evaluations. Hence, considering that $\frac{\partial}{\partial t} N_e = 5 \times 10^9$ electrons/cm^2/s for all the values of h we must obviously obtain an exaggerated evaluation of the error due to disregarding $\Delta\Phi_{int}$.

As shown in [4], determination of the phase difference at the receiver phasemeter input of the dispersion interferometer is performed from the record of what is termed the interference frequency i.e. from the curve for which the distance between the minima corresponds to a variation in phase difference at the input by 2π. Near the point $v_h \approx 0$ ($h \approx h_{max}$) the oscillation frequency of this curve corresponding to the $\frac{\partial}{\partial t}$ value $N_e = 5 \times 10^9$ electrons/cm^2/s is equal to $(F_i)\, h_{max} \approx 2{\cdot}5$ c/s.

The relative error in determining n_e from equation (5), due to disregarding $\Delta\Phi_{int}$ can be determined for each section of the rocket trajectory (on the above assumption regarding constant $\frac{\partial}{\partial t} N_e$ in the outer ionosphere at all altitudes) as

$$\frac{\delta n_e}{n_e} = \frac{(F_i)\, h_{max}}{F_i} \tag{8}$$

where F_i is the interference frequency recorded at the interferometer output corresponding to the given section of the trajectory.

Figure 5 shows the variation with altitude of the frequency F_i for the tests of 21st February and 27th August 1958. It is seen from these curves that this error is less than 10 per cent throughout excluding the section at the peak of the trajectory with a range of approximately 5 km. For this reason in the n_e versus altitude h curves drawn up in the present paper, data for altitudes near to the maximum reached in the flight are not employed, just as they were not previously in [1] and [2]. Data relating to the zone of maximum altitude attained are also omitted from the American n_e measurements

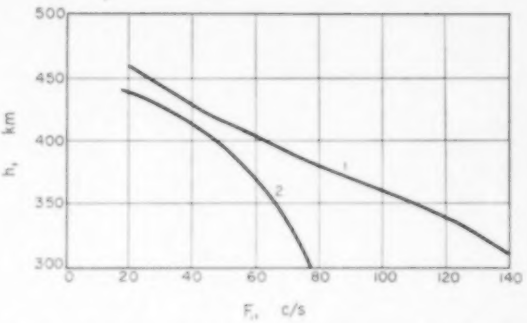


Fig. 5. Variation of interference frequency f_i for signals with frequencies of 144 and 148 Mc as a function of the height.
1—21st February 1958; 2—27th August 1958.

carried out by the dispersion method in the published results[1,2]. It should again be observed that the estimates of the effects of disregarding $\Delta\Phi_{int}$ on the n_e measurement results, based on the assumption of invariable $\frac{\partial}{\partial t} N_e$ values at all altitudes, are positively exaggerated.

Previously (see section 2 and commencement of the present section) concepts were introduced forming the basis of the permissible utilization of a number of approximate relations in calculating the preliminary results of the measurements, in particular for the coefficient of refraction of radio-waves in formula (2) the possibility of substituting increases in slant range ΔL by increases in altitude Δh and disregarding the variations in N_e during the flight time (for the whole trajectory apart from the zone near to the peak). We will now deal with errors due to the measurement equipment employed in the experiments. It follows from (5) (considering that $\Delta h = v_h \times \Delta t_1$):

$$\frac{\delta n_e}{n_e} = \frac{\delta(\Delta\Phi)}{\Delta\Phi} + \frac{\delta(\Delta t_1)}{\Delta t_1} + \frac{\delta v_h}{v_h} \tag{9}$$

Here $\frac{\delta n_e}{n_e}$ is the relative measurement error in n_e due to errors in determining $\Delta\Phi$, Δt_1 and v_h whilst $\frac{\delta(\Delta\Phi)}{\Delta\Phi}$, $\frac{\delta(\Delta t_1)}{\Delta t_1}$ and $\frac{\delta v_h}{v_h}$ are the relative measurement errors for the respective values.

An error occurs in determining n_e due to the variation in phase shifts occurring in the high-frequency channels of the instruments during the time Δt_1. These errors can attain high values with sudden voltage variation of the received signals at the receiver inputs (U_{in}). However, by leaving out of the treatment relatively small sections of the records where a sharp variation in U_{in} occurs, the relative error in determining n_e due to phase shifts in the instruments can be kept small. As seen from the specimen records [Fig. 7 on p. 42 of *Iskusstvennye Sputniki Zemli* No. 6], simultaneously with the recordings of the phase difference variations $\Delta\Phi$ the U_{in} variations are recorded on the same film. Employing the curve Fig. 11 in the paper[4], a study of the U_{in} variations on the sections of the records on which $\Delta\Phi$ is determined shows that the error in determining $\Delta\Phi$ due to variation in the signal amplitudes does not exceed 1·5 per cent.

It follows from the same paper[4] that the error in determining $\Delta\Phi$ due to variation in the signal frequency (on account of instability of the heterodynes and the Doppler effect) can be disregarded.

The error in determining Δt_1 is the sum of the recording instrument error and errors in working out the records. The recording instrument error is connected with irregular speed of travel of the film and enters in the determination of Δt_1 (where the interval Δt_1 is not a multiple of the interval between time marks) by interpolation carried out on the assumption of constant film speed. With $\Delta t_1 \approx 0{\cdot}2$ sec it can be considered that the error due to irregular film take-up is ± 2 per cent. The relative error falls off with increasing averaging time but increase in Δt_1 gives rise to an increase in the averaging interval for altitude Δh_1. In order to avoid a marked increase in the error due to irregular film take-up with a reduced averaging time and at the same time to avoid a marked increase in the averaging intervals with altitude, one and the same compromise value $\Delta t_1 = 0{\cdot}2$ sec was taken for all altitudes. For $\Delta t_1 = 0{\cdot}2$ sec, with maximum ascent height of the rocket $h_{max} = 500$ km at altitudes $h = 100$, 300 and 450 km the averaging intervals Δh_1 are approximately 500, 400 and 200 m respectively.

It can be seen from these specimen records that Δt_1 can suitably be calculated from the minima of the interference curve. With the averaging time selected, treatment of the records consists in determining a time interval Δt_1 approaching 0·2 sec during the course of which the phase difference variation $\Delta\Phi = 2\pi q$ (where q is a whole number) and in evaluating the number of cycles (q) of the interference curve on the section Δt_1. The time interval reading error on the film can amount to 0·002 sec; hence for $\Delta t_1 \approx 0{\cdot}2$ sec, the error in determining Δt_1 is approximately 1 per cent.

The total error in determining Δt_1 due to irregular film take-up and errors in treating the records is thus approximately 3 per cent.

The value $\frac{\delta v_h}{v_h}$ increases with diminishing v_h (i.e. on approaching the peak of the rocket trajectory). On the section of the trajectory used for determining n_e, $\frac{\delta v_h}{v_h}$ nowhere exceeds 1 per cent. Employing the discrepancy evaluations carried out above, it can be calculated from equation (9) that the relative error in determining n_e due to the measurement procedures does not exceed 6 per cent. The accuracy evaluations carried out above were made for n_e measurements by the dispersion interferometer method. When determining n_e from observations of rotation of the polarization plane the concepts regarding the size of the errors due to instability of the integral concentration N_e disregarding the effect of electron collision and utilizing altitude increments instead of slant range increments in the calculations continue in force. However since in this case $n_e(h)$ is averaged over considerably larger altitude intervals than in the radio-wave dispersion measurements the high precision determination of n_e lacks sense and therefore we will not dwell on the more detailed study of the precise determination of n_e from the Faraday effect.

It should be said in conclusion that the rate of variation of the integral electron concentration in the air column up to the peak of the

trajectory $\frac{\partial}{\partial t} N_e = 5 \times 10^9$ electrons/cm²/s is determined with adequate reliability.

The only cause which may give rise to an inaccurate calculation of the phase difference variations in proximity to the peak of the rocket trajectory could be the inaccurate determination of the moment when the rocket reaches the maximum altitude; however in order that the phase shift recorded in the experiment might be explained by the vertical motion of the rocket which may occur at this time, the error in determining the time corresponding to the trajectory peak would have to exceed the possible errors of the measuring devices employed.

6. EVALUATION OF RESULTS

The measurement results set out above afforded some material conclusions regarding the structure of the ionosphere. A comparison of the sections of the n_e versus h curves for altitudes above the principal ionization maximum (above 300 km) during the time of the three rocket launchings described shows that the rate of electron concentration decay with altitude (for $h_V \approx 300$ km) over one and the same geographical point can vary over a wide range (apparently as a function of time of year and day). This conclusion in particular contradicts the recently asserted opinion[13] regarding the existence of some model of the outer ionosphere only slightly affected by time of day, season, latitude, etc.

The inconstancy in time of the rate of decay with altitude of charged particle concentration in the upper ionosphere is confirmed by results obtained with ion traps installed on the third Soviet earth satellite[5].

Measurement results at the time of all three launchings fully confirm that the ionosphere has one principal electron concentration maximum at altitudes around 300 km and that the concept of a sharply defined so-called E layer of the ionosphere arose only as the result of the limited possibility of establishing a pure altitude electron concentration distribution on the basis of altitude-frequency characteristics obtained from ionosphere stations. This also followed from experiments carried out earlier[1, 12] during rocket launchings to altitudes up to 300 km.

It is necessary particularly to take up the determination of rate of integral electron concentration variation in the air column at altitude $h \approx 450$ km. Up to the present the literature reveals only evaluations of the value $\frac{\partial}{\partial t} N_e$ relating to the entire ionosphere layer.

These are based on radio-astronomical observations and observations of the Faraday effect with moon radar and yielded lower values than determined from the measurements described. In Ratcliffe's review paper[14], dealing with irregularities and horizontal motions in the ionosphere, values are derived $\frac{\partial}{\partial t} N_e \approx 10^8$-$10^9$ electrons/cm²/s. In Evans' paper[15] curves of integral electron concentration in the vertical column as a function of time are given, as illustrated in Fig. 6 (obtained with moon

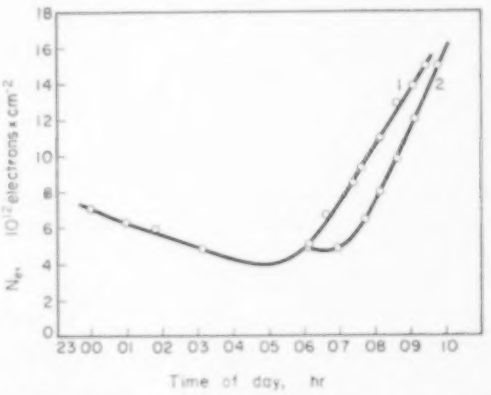


Fig. 6. Integral electron concentration as a function of time (obtained from observation of the Faraday effect with moon radar[15]).
1—October 1955; 2—November 1955.

radar). It is possible from these curves, obtained in 1955, to determine the maximum value (for these measuremets) $\frac{\partial}{\partial t} N_e \approx 10^9$ electrons/cm²/s. According to the data of V. V. Vitkevich and Yu. L. Kokurin[16], having observed rapid electron concentration fluctua-

tions, the rate of integral concentration variation $\frac{\partial}{\partial t} N_e \approx 10^8$ electrons/cm²/s.

It is interesting to assess the possibility of using the dispersion interferometer on rockets ascending to an altitude in the order of thousands of kilometres and also on artificial earth satellites. It follows from (4) that for determining the mean local concentration n_e on the section of the rocket path ΔL it is necessary to satisfy the condition $\Delta\Phi_{loc} > \Delta\Phi_{int}$, i.e.

$$n_e \frac{\Delta L}{\Delta t_i} = n_e v > \int_0^L \frac{\partial}{\partial t} n_e(l)\, dl, \qquad (10)$$

where v is the mean radial velocity of the rocket during the time Δt_i.

If $\int_0^L \frac{\partial}{\partial t} n_e(l)\, dl = \frac{\partial N_e}{\partial t} = 5 \times 10^9$ electrons/cm²/s, then, with $v = 10^6$ cm/s the reliably observed n_e concentration will be greater than several thousands electrons/cm³. Since at altitudes above 2,000 km, $n_e \approx 10^3$ electrons . cm³ and at altitudes $h > 25{,}000$ km, $n_e < 10^2$ electrons/cm³[17], then obviously electron concentration determination by the dispersion interferometer method is unreliable at distances of several thousands of kilometres from the earth's surface (due to variations in the integral electron concentration in the ionosphere) and are completely impossible at distances over 20×10^3 km. With the rocket at considerable distances from the earth there is a further cause giving rise to continuous variation in the integral concentration along the path from the observation point to the rocket. This cause is the variation in the path traversed in the ionosphere by the radio-waves emitted from the rocket to the receiving station, due to the diurnal motion of the earth. The concepts propounded lead on to the conclusion that the possibility of successfully studying local electron concentration in a medium by means of coherent radio-waves emitted from rockets is restricted to the zone of relatively high electron concentrations and small distances from the earth ($h < 1500$-2000 km). By installing a dispersion interferometer on a satellite it is possible with some tolerances to determine N_e and the fluctuations in N_e along the path from the satellite to the observer, but it is not possible to obtain reliable measuremens of n_e[18].

Sincere thanks are accorded to S. M. Rytov who rendered valuable advice at all stages of the work described in the present paper over a number of years.

REFERENCES

1. K. I. Gringauz, *Dokl. Akad. Nauk. SSSR.* **120,** 1234 (1958); *Isk. Sput. Zemli.* **1,** 62 (1958).
2. K. I. Gringauz and V. A. Rudakov, *Dokl. Akad. Nauk. SSSR.* **132,** 1311 (1960).
3. H. Friedman, Proc. IRE, **47,** 272 (1959).
4. K. I. Gringauz, V. A. Rudakov and A. V. Kaporskii, *Isk. Sput. Zemli,* present issue, p. 33.
5. K. I. Gringauz, V. V. Bezrukikh and V. D. Ozerov, *Isk. Sput. Zemli,* present issue, p. 63.
6. *Recent Investigations of Radiowave Propagation along the Earth's Surface,* (In Russian). (Ed. L. I. Mandel'shtam and N. D. Papaleksi), Gostekhizdat, Moscow, Leningrad (1945).
7. N. D. Papaleksi, *Transactions of the Expedition for Observation of the Total Solar Eclipse* 19.4.1936. (In Russian). Akad. Nauk. SSSR. Moscow. **1,** 115 (1937).
8. A. N. Shchukin, *Physical Bases of Radiowave Propagation in the Ionosphere,* Svyaz'izdat Moscow (1940).
9. N. S. Gerzon, *UFN.* **42,** 546 (1952).
10. J. C. Seddon and J. E. Jackson, *IGY World Date Centre A, Experimental Results of the U.S. Rocket Programmes for the IGY to 1 July* 1958, National Academy of Sciences. Washington. **1,** 140 (1958).
11. J. C. Seddon, *J. Geoph. Res.,* **59,** 463 (1954).
12. J. E. Jackson, *J. Geoph. Res.* **59,** 377 (1954).
13. Ya. L. Al'pert, E. F. Chudesenko and B. S. Shapiro, *Isk. Sput. Zemli.* No. 1 (1958).
14. J. A. Ratcliffe, *The Physics of the Ionosphere,* Physical Soc., London (1956).
15. J. V. Evans, *Proc. Phys. Soc.,* **69B,** 953 (1956).
16. V. V. Vitkevich and Yu. L. Kokurin, *Radiotekhnika i Elektroika,* **3,** 1373 (1958).
17. K. I. Gringauz, V. G. Kurt, V. I. Moroz and I. S. Shklovskii, *Dokl. Akad. Nauk. SSSR.* **132,** 1062 (1960). *Isk. Sput. Zemli.* Present issue p. 108.
18. K. I. Gringauz, *UFN.* **69,** 345 (1959).

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 194–196. Printed in Great Britain.

ABSTRACTS

No. 7 (1961)

*E. A. Benediktov, G. G. Getmantsev and V. L. Ginsburg: **Radioastronomical investigations employing artificial satellites and space rockets** (pp. 3–22).

One reason for the efficacy of radioastronomical methods is the possibility of using a very wide range of frequencies, extending over approximately eleven octaves as against the range, inside of one octave, corresponding to the visible spectrum. The only radical method of further extending the frequency range involves artificial earth satellites and space rockets reaching out beyond the earth's atmosphere. The present paper emphasizes basic measurement principles. Space rockets in the vicinity of Mars, Venus and other planets of the solar system can provide radio emission measurements over a wide range of frequencies. The ground antenna for measuring the radio emission on a one-metre wavelength from Mars and Venus would require a 150 m diameter mirror. On the other hand the dimensions for an antenna on a space rocket would approach the wavelength, depending on the solid angle presented by the planet to the space rocket. A study of cosmic radio emission and the radio emission of discrete sources yields a table illustrating radiowave absorption under cosmic conditions for wavelengths 0·3 km, 1 km and 3 km. Corpuscular solar fluxes can have an appreciable influence on wavelengths above 200 to 300 m. From the standpoint of instrumentation, simple antennae are adequate for artificial satellites, employing the travelling-wave type for directional reception from discrete sources against a substantial dispersed cosmic radio-emission. Artificial satellite measurements based on diffraction of extra-terrestrial radio-emission at the edge of the earth's and the moon's disks are of interest in the angular localization of discrete sources and for defining details in the distribution of nonthermal cosmic radio-emission on the celestial vault. Artificial satellite experiments are indicated for establishing the theoretical basis for the radio-emission in the radiation belts which are sufficiently intense for measurement. From the standpoint of electron-concentration measurements in the ionosphere and in interplanetary space, the most promising may be the method of measuring the group delay time of artificial satellite transmitter signals synchronized with low-frequency modulated oscillations. The magnetic field has a negligible effect on radiowave propagation on a wavelength of 600 m at altitudes above 3000 km. The survey draws on a bibliography of 39 references.

V. V. Vitkevich, A. D. Kuz'min, R. L. Sorochenko and V. A. Udal'tsov: **Results of radioastronomical tracking of Soviet space rockets** (pp. 23–31).

The radio-interference method widely employed in radio-astronomy for determining the angular coordinates of discrete sources of radio-emission was used for tracking the first, second and third Soviet space rockets. The observations were carried out on a frequency of 183 Mc/s. For determining the intensity of a signal from the space rocket use was made of the emission from a discrete source in the Cygnus constellation with an energy flux density of 70×10^{-24} W . m^{-2} . c/s on the 1·5 m wavelength. Time characteristics of slow signal-intensity variations, referred to an isotropic emitter at the distance of the space rocket, show that in addition to the rapid variations there were slower variations from the first rocket with characteristic periods, one of 8 to 12 min and the second from 40 min to one hour. On the second space rocket, variations in signal intensity occurred with a period of 45 min, falling to 10-13 min. An analysis of records from the third space rocket employing two mutually perpendicular antennae showed fading from maximum to minimum over a period of approximately 3 min, associated with rotation. Slower variations in the mean amplitude of the signal with periods greater than 20-30 min are attributed to the Faraday effect.

V. V. Beletskii and Yu. V. Zonov: **Rotation and orientation of the third Soviet satellite** (pp. 32–55).

A method is set out for determining the parameters of the satellite motion around the mass centre and its orientation in space from the angles of rotation of the magnetometer frame relative to the body of the satellite as measured by two pickups and telemetered to earth. Results from the first 109 loops of the satellite orbit demonstrate the possibility of calculating the orientation of any instrument on the satellite. Precession and rotation periods were determined with an accuracy of 5 sec, the angle between the satellite axis and precession axis, coinciding with the direction of the kinetic moment vector, i.e. angle of nutation, with an accuracy of 1°, and the position of the precession axis with an accuracy of 10°. The satellite rotated around the direction of the kinetic moment vector in such a way that the angle between the satellite axis and the kinetic moment vector was in the neighbourhood of 90°, deviating from this position on different loops by not more than 6°. The precession period increased slowly from 135-140 sec in the early loops (1-5) to 195 sec on the 283rd loop. The slow variation in the parameters of satellite rotation and orientation may be due to the effect of perturbing factors such as the influence of aerodynamic and gravitational perturbations, the interaction of the satellite current systems with the earth's magnetic field and, to a lesser degree, Foucault currents, etc.

*A. D. DANILOV: **The formation of O_2^+ ions in the upper atmosphere** (pp. 56–59).

The fact that O_2^+ ions are observed up to altitudes of 400 km, whilst molecular oxygen must be completely dissociated above 200-250 km gives rise to doubts that the capture and loss mechanism is responsible for the formation of O_2^+ ions at all altitudes where these ions occur. A quantitative study yields the impression that an explanation of the presence of O_2^+ ions at high altitudes has to be sought from a formation mechanism independent of the O_2 molecule concentration. It is found that the rate of formation of O_2^+ ions in accordance with the recombination reaction of neutral and ionized oxygen atoms closely coincides with the rate of disappearance of these ions. The proposed mechanism explains the daily variations in ion concentration. At altitudes 160-180 km the relative O_2^+ concentration (fraction of O_2^+ ions in the total quantity of ions) increases approximately four times during the transition from evening to morning, whilst due to the evening decay in overall electron density the absolute O_2^+ ion concentration diminishes by a factor of 2 to 3.

S. P. YATSENKO: **The ionization of gases entrained by a satellite in the upper atmospheric layers** (pp. 60–63).

In the first hours of existence of the third Soviet artificial satellite, the radio frequency mass spectrometer recorded ions with mass number 18 not associated with the ionosphere. These ions originated from H_2O^+ ions. An attempt is made to find the cause for the ionization of water vapour desorbed from the satellite surface, on the basis of purely qualitative concepts, from the shape of the outgassing curve. A scattered molecule flux curve is obtained through the retarding field of the instrument as a function of time. The position of the maxima on the scatter curve is determined with the same accuracy as the satellite orientation, the possible time discrepancy being ±10 sec. It is shown that the maxima of the outgassing curve and of the scatter curve coincide in time. Consequently the apparatus records basically molecules subjected to collision with atmospheric particles. The energy of encountered particles in respect to the satellite is 5-10 eV, being inadequate for impact ionization. The outgassing-ion current varies within narrower limits, conforming to outgassing-molecule scatter by atmospheric ions. It is concluded that capture and loss is the most probable cause of ionization.

*V. G. ISTOMIN: **Investigation of the composition of the Earth's atmosphere on geophysical rockets 1957–1959** (pp. 64–77).

Four experiments were conducted during the period, for positive-ion mass spectrum determination at altitudes 90 to 210 km, using the ion radiofrequency mass spectrometer. Measurements were made in the middle latitudes of the European U.S.S.R. in the morning, with height of sun 0° (two launchings) and 36°, and in the evening, with sun sinking to 6°. All four experiments, carried out at different times, under different conditions and employing different instruments, provided the same conclusion regarding the relative composition of the atmosphere. Hydrogen and helium ions are not observed at 100-200 km. At different moments of local solar time the relative O^+ ion concentration in the 150-200 km range increases with increasing local solar time. The longitude effect, or possible variations in the composition of the ionosphere, may only mask the latitude effect, which in fact may be greater. Ionization of gas evolved by the rocket proceeds as the result of charge exchange with atmospheric ions.

*V. G. KURT and V. I. MOROZ: **The potential of a metal sphere in interplanetary space** (pp. 78–88).

An attempt is made to assess to a first approximation the potential of the container and instrumentation employed on space rockets, considering a roughly spherical container with a radius of 65 cm. The effect of motion and the magnetic field, being 1-2 V, can in this case be disregarded in the light of the indeterminacy of other factors. Variations in diameter by some multiple and slight deviations in shape are also not substantial. In the illuminated section of the trajectory the container potential is between $-2{\cdot}5$ V and $+4$ V for a concentration of 10 particles cm^{-3}. In the outer belt on the illuminated section of the trajectory the radiation electrons do not in the main affect the potential, at any rate in periods of magnetic quiescence. However, high negative potentials (some kilovolts) are not ruled out if the standard concepts of very high concentrations of soft radiation electrons in the outer zone are correct, and if the interplanetary gas in the outer zone is sufficiently rarefied.

*A. A. POKHUNKOV: **Mass-spectrometry investigations of the structural parameters of the earth's atmosphere at altitudes between 100 and 210 km** (pp. 89–100).

The measurement of absolute and relative concentrations of the separate gases at different altitudes, together with the mean molecular weight, yields information on the presence and level of gravitational separation, quantitative characteristics of the dissociation processes and important data regarding the composition of light gases, necessary for solving problems connected with dissociation of the gases of the earth's atmosphere. The mass-spectrometer employed consisted of an analyzer with ion source, pre-amplifier and measurement unit. The ion source had a flanged joint allowing removal of the ion source for utilizing the instrument for analyzing the ion composition of the atmosphere. A known gas mixture of H_2, He, Ne and Ar was carried for

checking performance of the instrument and calibration. Two experiments were carried out in July 1959 for investigating the neutral composition of the outer atmospheric layers. For true readings at high altitude the container was detached from the rocket and analysis proceeded at some hundreds of metres from the rocket. In the second experiment gases were recorded having mass numbers 14, 16, 18, 28, 32, 40 and 44. In the H_2O, OH and H_2 and H_1 group the H_2O ion peak showed the greatest amplitude, being independent of the position of the analyzer axis relative to the velocity vector. In the case of the N_1, O_1, N_2, O_2, Ar, CO_2 and N_2O group, in addition to the general decay with altitude, there occurred periodic variations with time of flight. The mean molecular weight of gases inside the spectrometer varies very slightly with altitude. A mean molecular weight of 26·6 g/mol is suggested at 200 km.

*A. V. GUREVICH: **Perturbations in the ionosphere due to travelling bodies** (pp. 101–124).

The special nature of motion in the outer atmosphere resides primarily in the fact that it proceeds in an extremely rarefied medium, and the particle range is much greater than the size of the body. Consequently the usual hydrodynamic methods of investigating the interaction of a moving body with the medium are no longer valid. It is necessary to employ the kinetic theory, considering the gas not as a continuous medium but as an aggregation of separate particles. Secondly we are dealing with the motion of a body in a plasma, and the interaction of the body with ions and electrons has to be considered. The different effects produce different perturbations in the electron and ion concentration. The quasi-neutral state of the plasma is destroyed and an electric field is produced. In calculating the values of the perturbations it is assumed throughout that the speed of bodies in the ionosphere is much greater than the thermal velocity of the molecules or ions and much less than the thermal velocity of the electrons. Section 1 is given over to the calculation of neutral particle concentration near the moving body. In section 2 the perturbations of electron and ion concentrations is sought; the value of the electric field is determined. Section 3 resolves the same problem taking into account the earth's magnetic field.

*YA. L. AL'PERT: **The study of the ionosphere and interplanetary gas employing artificial earth satellites and space rockets** (pp. 125–169).

Experimental possibilities are examined in relation to the outer section of the ionosphere extending 300 to 400 km beyond the principal maximum and several thousand kilometers from the earth, passing over into interplanetary gas, having in view the study of charged particle density, elastic collisions, disuniform electron or ion formations, plasma waves, and the electromagnetic emission of particles. Since the instruments are housed in a body travelling at a speed exceeding or comparable with the thermal velocity of the particles of the gaseous medium it is necessary to take into account the interaction of the body with the plasma and with the radiation field in which it is travelling. These difficulties are enhanced by the intervening ionosphere zones introducing masking and perturbing effects. A brief account of results to date points to the necessity for a diversity of methods in line with the diversity of the problems. One group of experiments is based on the analysis of radiowaves emitted from the satellite and received at different points on the ground. A second group employs probes, transmitting the readings to the earth. A bibliography of 49 references is drawn upon for the purpose of assessing the problems involved mainly employing radio methods. Interest derives also from the study of plasma as such, independently of the various geophysical or cosmic-physics problems.

Planet. Space Sci. Pergamon Press 1961. Vol. 8, pp. 197–201. Printed in Great Britain.

RESEARCH NOTES

On the night sky hydroxyl emission

(*Received* 10 *August* 1961)

At present the incompatibility of data on the hydroxyl emission intensity and atomic hydrogen concentration below 100 km has become quite evident. Number $\mathcal{N}$ of newly formed vibrationally excited hydroxyl molecules over 1 cm^2 of the earth's surface can be represented as

$$\mathcal{N}_1 \sim h \,.\, [O_3] \,.\, [H] \,.\, \alpha_1$$

in the case of an ozone-hydrogen reaction[1] and

$$\mathcal{N}_2 \sim h \,.\, [O_2^*] \,.\, [H] \,.\, \alpha_2$$

in the case of a reaction of vibrationally excited oxygen molecules with hydrogen atoms[2,3,4]. Here $[O_3]$ and $[O_2^*]$ are ozone and vibrationally excited oxygen molecules concentrations, respectively, [H] is atomic hydrogen concentration, α_1 is the rate coefficient of the ozone-hydrogen reaction (according to Bates and Nicolet[1] $\alpha_1 \sim 10^{-12}$ cm^3/sec), α_2 is the rate coefficient of the reaction of vibrationally excited oxygen molecules with hydrogen atoms (according to our evaluations[4] $\alpha_2 \sim 10^{-10}$ cm^3/sec) and h is the scale height ($h \sim 10^6$cm). According to Friedman's latest data[5] atomic hydrogen concentration at an altitude somewhat lower than 100 km is hardly more than $5 \cdot 10^6$ cm^{-3} and at lower heights it cannot be greater since in the dense layers of the atmosphere atomic hydrogen rapidly forms chemical compounds and cannot be in a free state for a long time. In the light of the above considerations, since $\mathcal{N} \sim 10^{12}$ of vibrationally excited hydroxyl molecules cm^{-2}/sec[4], we get $[O_3] \sim 2 \cdot 10^{11}$ cm^{-3} and $[O_2^*] \sim 2 \cdot 10^9$ cm^{-3}. Such enormous $[O_3]$ and $[O_2^*]$ values seem to be unacceptable. If Friedman's data will not be refuted, one should seek for new mechanisms of formation of vibrationally excited hydroxyl molecules.

In this case the only way out seems to be a supposition about the domination in the upper atmosphere at altitudes lower than 100 km of such atomic hydrogen compounds (not of atomic hydrogen itself) which are capable of entering chemical reactions with the yield of vibrationally excited hydroxyl molecules. Hydrogen, hydroxyl and perhydroxyl molecules are apparently such compounds. Here is the list of possible reactions:

$$O_2^* + H_2 \longrightarrow OH^* + OH^*,$$
$$O_2^* + OH \longrightarrow OH^* + O_2,$$
$$O_2^* + HO_2 \longrightarrow OH^* + O_3,$$
$$O_3 + HO_2 \longrightarrow OH^* + O_2 + O_2.$$

Asterisks denote molecules in a vibrationally excited ground state. The existence of free hydroxyl is possible only where atomic oxygen concentration is not great and, therefore, the appearance of atomic oxygen as a result of the reaction

$$OH + O \longrightarrow O_2 + H$$

is hardly probable.

According to the latest data[6] the reaction

$$H + O_2 + M \longrightarrow HO_2 + M$$

is by three orders of magnitude more effective than the reaction

$$O + O_2 + M \longrightarrow O_3 + M$$

where M is any molecule or atom of the upper atmosphere. Therefore, there are essential causes of rapid disappearance of atomic hydrogen below the 100 km level. At high altitudes atomic hydrogen can disappear due to diffusion upwards and dissipation*.

Thus excited hydroxyl can be formed as a product of various chemical reactions in the upper atmosphere. Each of them will ensure its own peculiar relative excitation of hydroxyl vibrational levels. Intensity and height variations of various reactions will lead to variations

* It is also possible that perhydroxyl is formed in the reactions

of relative population rates and rotational temperatures at these levels. This is precisely what is found[7]. Further studies will elucidate the nature of the hydroxyl emission.

$$OH + O_3 \rightarrow HO_2 + O_2,$$
$$H + O_3 \rightarrow HO_2 + O_2.$$

REFERENCES

1. D. R. Bates and M. Nicolet, *J. Geophys. Res.*, **55**, No. 3, 301 (1950).
2. V. I. Krasovskii and V. T. Lukashenya, *Dokl. Akad. Nauk U.S.S.R.*, **80**, 735 (1951).
3. V. I. Krasovskii, *Airglow and Aurorae* (ed. Armstrong and Dalgarno), Pergamon Press, London, p. 193 (1956).
4. V. I. Krasovskii, *Adv. Phys. Sci., Moscow*, **63**, No. 4, 673 (1957).
5. H. Friedman, Proc. XI Internat. Astronautical Congress, Stockholm, 83 (1960).
6. F. Kaufman, *Proc. Roy. Soc.* **247 A**, No. 1248, 123 (1958).
7. V. I. Krasovskii, N. N. Shefov and V. I. Yarin, *J. Atmos. Terr. Phys.*, **21**, 46 (1961).

V. I. Krasovskii

The Institute of Physics of the Atmosphere, The U.S.S.R. Academy of Sciences, Bol. Cruzinskaja 10, Moscow

On the excitation of subvisual O1 λλ6300-6364 arcs at middle latitudes

(*Received* 16 *October* 1961)

Subvisual OI λλ 6300-6364 arcs at middle latitudes have been investigated photometrically by Barbier[1,2], Roach and Marovich[3,4], and Duncan[5]. The phenomenon appears to be not uncommon, being associated with large geomagnetic storms.

Barbier[1] has summarized the morphology as follows: "At the time of a very strong auroral perturbation a red arc is probably formed in the zone of the maximum activity and shifts toward the south; it is then checked and becomes stationary at a latitude which is slightly higher than 45° (geomagnetic), then later it disappears slowly in place without spreading and remaining unaffected by later recurrence of auroral activity." From reports by the various investigators the observational characteristics may be listed:

(a) They occur at times of general strong auroral and magnetic activity on the side nearest to the equator of the main aurora. One observer reports two occasions on which the arc appeared on the night following a bright visual aurora.

(b) They are situated at an altitude of 350 to 450 km; their width in latitude is of the order of 500 km; and their orientation is reported by different observers to be along magnetic dip latitude and parallel to theoretically computed isochasms.

(c) All observers note the persistence and stability of the luminosity; the brightness always decreases from the beginning to the end of the observations and the decay of the luminosity is remarkably constant for each arc (see Fig. 1)

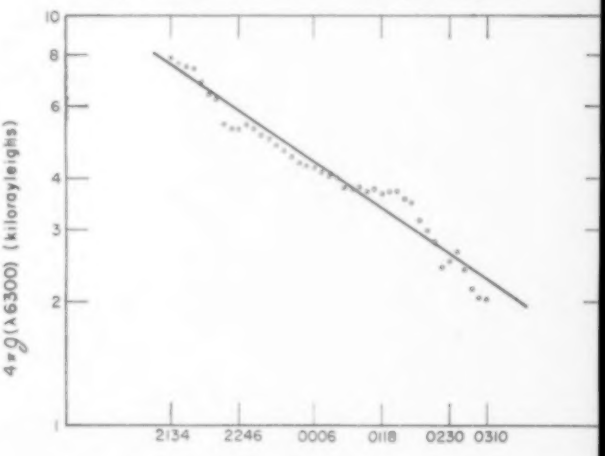


Fig. 1. Zenith photon emission rate of a stable red arc as a function of time recorded by F. E. Roach and E. Marovich[3] at Fritz Peak, Colorado on September 29/30, 1957.

and from one to the next. The decay rates for several stable red arcs are listed in Table 1.

(d) The absolute intensity of λ6300 in the arc, $4\pi\mathcal{J}$ (6300), ranges from a few hundred rayleighs to as much as 13 kilorayleighs. The ratio of photon emission rates $4\pi\mathcal{J}$ (6300)/$4\pi\mathcal{J}$ (5577) is reported to be 80:1 or greater. Barbier[2] states, from photometric evidence, that

no N_2^+ IN emission is associated directly with the arc.

Compared with the intensity fluctuations usually observed during an auroral display reaching local zenith and lasting an entire night the steady decay of the stable red arc is remarkable. The numerical values of the decay constant closely correspond to the electron loss coefficient in the $F2$ region[6] of the ionosphere suggesting the red arc to be a consequence of normal ionospheric recombination processes enhanced by the auroral disturbance.

Recombination in the $F2$ region occurs through the two step process of ion-atom exchange between the dominant O^+ ions and the molecular nitrogen and oxygen, followed by dissociative recombination of the molecular ions with the effective recombination rate governed by the slower ion-atom exchange reaction. The oxygen atoms produced may be left in the upper levels of the transitions leading to the $\lambda\lambda$ 6300-6364 lines. It is now believed that the sequence accounts for the presence of these lines in the twilight and nightglow[7].

According to one hypothesis[8] the 100 fold, or greater, photon emission rate in the red arc over the average nightglow is attributed to increased recombination due to an increase in the concentration of molecular nitrogen, a 400 fold increase being required. Even if it were possible to maintain such a high localized concentration at 400 km the effective recombination rate required by the hypothesis would now be some 2×10^{-3} sec^{-1} whereas the observed decay of the red arcs is about 7×10^{-5} sec^{-1} (see Table 1).

Table 1

Date	*Reference*	*Maximum Intensity* $4\pi\mathcal{J}$ *Kilorayleighs*	*Decay Constant* sec^{-1}
23 - IX - 57	Barbier[2]	4	$6{\cdot}2 \times 10^{-5}$
29 - IX - 57	Barbier[1]	13	$1{\cdot}1 \times 10^{-4}$
29 - IX - 57	Roach[3]	8	$6{\cdot}0 \times 10^{-5}$
11 - II - 58	Barbier[2]	6·4	$7{\cdot}8 \times 10^{-5}$
5 - IX - 58	Duncan[5]	7	$4{\cdot}0 \times 10^{-5}$
22 - X - 58	Roach[4]	6	$4{\cdot}2 \times 10^{-5}$
28 - XI - 59	King & Roach[8]	4	$1{\cdot}1 \times 10^{-4}$
5 - XII - 59	Barbier[2]	1·5	$4{\cdot}1 \times 10^{-5}$

Absence of the N_2^+ λ 3914 emission in the stable arc[2] suggests that primary incident energetic particles are not responsible for the excitation.

The middle latitude red arc may be due entirely to a current flowing in the high atmosphere. A field of a few volts per km would ensure a high ratio of $4\pi\mathcal{J}$ (6300)/$4\pi\mathcal{J}$ (5577) and an electron density of 10^4 to 10^5 cm^{-3} would suffice to account for the observed luminosity. Existence of the current throughout the night* would be required and the decay of the luminosity could be attributed to the usual night time decay of electron density. The stable arc persists even through periods of active aurora[2] though the emission heights may well be different.

The energy distribution of electrons under the influence of an electric field has been derived[9] in connection with a discharge mechanism of auroral excitation. The treatment, however, neglected inelastic collisions of energetic electrons with N_2 molecules since pertinent data were not available at the time. Recent laboratory measurements indicate that the cross section for excitation of vibrational levels in N_2 in the energy range 1·7 to 4 eV is high[10]. Therefore N_2 may be the principal constituent accounting for the energy limitation of electrons in a current stream up to an altitude at which its concentration becomes small. At 400 km N_2 amounts to about 5 to 50 per cent of the total particle concentration, depending on the model atmosphere adopted, while atomic oxygen is the principal constituent. Excitation of the ^1D levels of OI then becomes the principal loss mechanism for the energetic electrons. This may account for the altitude at which the stable red arc is observed.

Acknowledgment—This investigation was carried out, in part, at the Department of Applied Mathematics, Queen's University, Belfast. I have profited greatly from discussions with Professor D. R. Bates. I am

* There are no reports of the arc disappearing completely and suddenly before observations were normally terminated by encroaching twilight or cloudy conditions.

indebted to Dr. L. R. Megill for calling to my attention the importance of vibrational excitation of N_2 in the energy limitation of electrons. Financial support was obtained through U.S. Air Force Grant AF-EO ARDC 61-16 and National Science Foundation Grant G-14898.

REFERENCES

1. D. Barbier, *Ann. Géophys.* **14**, 334 (1958).
2. D. Barbier, *Ann. Géophys.* **16**, 544 (1960).
3. F. E. Roach and E. Marovich, *J. Res. Nat. Bur. Stand.* **63D**, 297 (1959).
4. F. E. Roach and E. Marovich, *J. Res. Nat. Bur. Stand.* **64D**, 205 (1960).
5. R. A. Duncan, *Austr. J. Phys.* **12**, 197 (1958).
6. J. A. Ratcliffe, E. R. Schmerling, C. S. Setty and J. O. Thomas, *Phil. Trans. Roy. Soc.* **A248**, 621 (1956).
7. J. W. Chamberlain, *Astrophys. J.* **127**, 54 (1958).
8. G. A. M. King and F. E. Roach, *J. Res. Nat. Bur. Stand.* **65D**, 129 (1961).
9. J. W. Chamberlain, *The Airglow and the Aurora*, (ed. E. B. Armstrong and A. Dalgarno) p. 206. The Pergamon Press, London (1955).
10. G. J. Schulz, *Phys. Rev.* **116**, 1141 (1959).

M. H. Rees

Department of Physics,
University of Colorado,
Boulder, Colorado.

Further comments on the origin of the *D*-region

(*Received* 28 *July* 1961)

Recently, it was suggested that photoionization of highly vibrationally excited ground state 0_2 molecules by solar Lyman-α may be an important process in the formation of the *D*-region of the ionosphere[1]. In an earlier publication a detailed analysis of the formation of the *D*-region was presented by Nicolet and Aikin[2] in which they showed that at about 80 km the dominant process was photoionization of NO by Lyman-α. However, it has since been shown[3] that the recombination coefficient of NO^+ assumed by them in their analysis was much too low and the higher measured recombination coefficient, when applied to their analysis, gave much lower electron densities at 80 km than is compatible with ionospheric observations. It is, therefore, proposed here to present further arguments which tend to support the previously suggested process of photoionization of vibrationally excited 0_2 molecules.

We have considered in detail only one of the two processes suggested previously for producing vibrationally excited ground state 0_2 molecules, namely, the purely optical process. Briefly, for example, this process involved absorption in the (2, 0) Schumann-Runge (*SR*) band followed by emission in the (2, $v''>11$) SR bands, thus producing $0_2\,(X^3\Sigma_g^-)_{v''>11}$ molecules. The assumption was then made that the latter molecules may be photoionized by Lyman-α radiation.

In the following calculations we have considered only the ionization at 80 km for an overhead quiet sun. The basic data for this altitude were:

(1) density of $0_2\,(X^3\Sigma_g^-)_{v''=0}$ $n_{0_2}=8{\cdot}93\times 10^{13}$ cm^{-3}

(2) intensity of solar radiation[4] at $\lambda\approx 1984$ Å, $I_{\Delta\lambda}=0{\cdot}1\,\mu$W cm^{-2} Å^{-1}

(3) intensity of solar Lyman-α at $\lambda=1216$Å, $I_{Ly\alpha}=2$ ergs cm^{-2} sec^{-1}

The rate of production of $0_2\,(B^3\Sigma_u^-)_{v''=2}$ molecules, q, by absorption in the (2, 0) SR bands was first calculated. The absorption coefficient for this band, measured by Ditchburn and Heddle[5], was used since there was a general agreement in the oscillator strength of the SR bands measured by them and a number of other reported results[6], although Bethke[7] reported a much lower value. Thus, the average absorption cross-section for the (2, 0) SR band, $\sigma_{(2,0)}=1{\cdot}64\times 10^{-21}$ cm^2, where the average was taken over a band width of $\Delta\lambda=25$Å (between band heads $\lambda=1971$ and $\lambda=1998$). Furthermore, within this $\Delta\lambda$ $I_{(2,0)}=2{\cdot}5\times 10^{12}$ quanta cm^{-2} sec^{-1}, the average solar radiation flux. Hence,

$$q=n_{0_2}\cdot\sigma_{(2,0)}\cdot I_{(2,0)}=3{\cdot}66\times 10^{5}\ \text{cm}^{-3}\ \text{sec}^{-1} \quad (1)$$

Clearly, by spontaneous emission these electronically excited 0_2 molecules will radiate and return to the ground state, where most of these will terminate at $v''>0$ because of the Franck-Condon principle. It was assumed that one-half of these end in $X^3\Sigma_g^-\,(v''\geqslant 11)$, the actual distribution depending on the relevant Franck-Condon factors. We therefore have for the

production rate of vibrationally excited ground state 0_2 molecules with $v'' \geqslant 11$,

$$q' = 1{\cdot}83 \times 10^5 \text{ cm}^{-3} \text{ sec}^{-1}. \quad (2)$$

The lifetime of these metastable vibrationally excited 0_2 molecules is determined by the radiative transition probability and collisional deactivation. The former has been estimated to be about 10^{-9} of the ordinary dipole radiation[8], hence a radiative lifetime of about 10 sec. On the other hand, collisional vibrational relaxation[9] may be as much as 100 secs at the density and temperature existing at 80 km. This would be the case for deactivation involving $v'' = 1$ to $v'' = 0$. As a lower limit an average lifetime of about 1 sec was assumed. Therefore, from (2) the concentration of vibrationally excited 0_2 ($v'' \geqslant 11$), $n' = 1{\cdot}83 \times 10^5$ cm^{-3}.

In order to calculate the rate of ionization, Q, of these vibrationally excited 0_2 molecules an estimate was made of the ionization cross-section, σ_i, for the following Franck-Condon transition[1]:

$$0_2\,(X^3\Sigma_g^-)_{v''=11} + h\nu\,(Ly\alpha \longrightarrow 0_2^+\,(X^2\Pi_g) + e^- \quad (3)$$

It was assumed that σ_i was not much different from that for the ionizing transitions arising from $v'' = 0$. Cross-section measurements of the first ionization continuum[10] have been reported for the latter, the average cross-section being about 2×10^{-18} cm^2. It should be noted that the spectral region of this ionization continuum is overlain with strong pre-ionized absorption bands so that, if process (3) lies within one of these bands σ_i may be considerably larger. However, we adopted the following as a reasonable estimate, $\sigma_i = 2 \times 10^{-18}$ cm^2. The rate of ionization is then given by,

$$Q = n' \cdot \sigma_i \cdot I_{Ly\alpha} = 0{\cdot}0446 \text{ cm}^{-3} \text{ sec}^{-1}. \quad (4)$$

It now remained to calculate the equilibrium electron density, n_e. Using the value of the recombination coefficient of 0_2^+ adopted by Nicolet and Aikin[2], namely $\alpha_{0_2} = 3 \times 10^{-8}$ cm^3 sec^{-1}, then

$$Q = \alpha n_e^2, \quad (5)$$

where we neglected the effect of negative ions since, at 80 km, their concentrations were apparently much lower than n_e. Substituting the values of α_{0_2} and Q in (5) we have,

$$n_e = 1{\cdot}2 \times 10^3 \text{ cm}^{-3}. \quad (6)$$

As a final remark, it should be noted, as discussed previously[1], that production of vibrationally excited 0_2 molecules by chemical reactions should also be considered. Reference is made to the reaction

$$0\,(^1D) + 0_3 \longrightarrow 0_2^* + 0_2. \quad (7)$$

If the rates of this reaction and others connected with it were sufficiently large to yield a production rate as large as (1) then the proposed mechanism discussed above can more than adequately account for the observed ionization in the D-region.

REFERENCES

1. E. C. Y. Inn, *Planet. Space Sci.* **5**, 70 (1961).
2. M. Nicolet and A. C. Aikin, *J. Geophys. Research* **65**, 1469 (1960).
3. E. C. Y. Inn (to be published).
4. H. H. Malitson, J. D. Purcell, R. Tousey and C. E. Moore, *Astrophys. J.* **132**, 746 (1960).
5. R. W. Ditchburn and D. W. O. Heddle, *Proc. Roy. Soc.* (*London*) **A226**, 509 (1954).
6. C. E. Treanor and W. H. Wurster, *J. Chem. Phys.* **32**, 758 (1960).
   J. C. Keck, J. C. Camm, B. Kivel and T. Wentink, Jr., *Ann. Phys.* **7**, 1 (1959).
7. G. W. Bethke, *J. Chem. Phys.* **31**, 669 (1959).
8. G. Herzberg, *Molecular Spectra and Molecular Structure I. Spectra of Diatomic Molecules* (2nd Ed.), Van Nostrand, New York, p. 131 (1950).
9. V. Blackman, *J. Fluid Mech.* **1**, 61 (1956).
10. Po Lee, *J. Opt. Soc. Amer.* **45**, 703 (1955).

E. C. Y. Inn

Lockheed Missiles and Space Company, Sunnyvale, California.